

STIC-Biot ch/Ch mLib

92268

From: Li, Ruixiang  
Sent: Wednesday, April 23, 2003 6:16 PM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search of Application NO: 09/898,586

Please do a standard search on SEQ ID NO: 23 against commercial nucleic acid databases.

Thank you very much!

Ruixiang Li  
GAU 1646  
CM1 10E18  
Mail Box 10D19  
306-0282

Point of Contact:  
Toby Port  
Technical Info. Specialist  
CM1 6A04  
703-308-2634

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 4/24  
Date Completed: 5/2  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 03:54:25 ; Search time 279 Seconds

(without alignments)  
8184.682 Million cell updates/sec

Title: US-09-898-586-23

Perfect score: 1014

Sequence: 1 taacactctctcctaacca,.....tatgcatgtgactgacaa 1014

Scoring table: IDENTITY\_NUC

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	991	97.7	1014	22	AAS09956	DNA encoding human
2	987.4	97.4	1012	22	AAS09955	DNA encoding human
3	983.6	97.0	1040	22	AAS09946	DNA encoding human
4	942	92.9	1315	22	ABA09073	Human olfactory re
5	936	92.3	974	22	AAF58613	Human RECAP polyu
6	905.4	89.3	930	22	AAH32366	Human olfactory re
7	903.6	89.1	933	24	ABK65141	CDNA encoding huma
8	900.6	88.8	930	22	AAH31617	Human olfactory re
9	900.6	88.8	930	22	AAH31648	Human olfactory re

C	10	545	53.7	1957	22	ABA45207	Human breast cell
C	11	545	53.7	1957	22	ABA55696	Human foetal liver
C	12	545	53.7	1957	22	AAI13969	Probe #3902 for ge
C	13	545	53.7	1957	22	AAI35352	Probe #4038 used t
C	14	545	53.7	1957	24	ABS03949	Human genome-deriv
C	15	538.2	53.1	2282	22	AAS08660	Human cDNA encodin
C	16	536.8	52.9	940	22	ABN89130	Human cDNA encodin
C	17	536.6	52.9	1788	22	AAS08655	Human cDNA encodin
C	18	536.6	52.9	1735	22	AAH32363	Human cDNA encodin
C	19	535.4	52.8	930	22	AAH32363	Human olfactory re
C	20	535.2	52.8	933	22	ABAS0343	Human breast cell
C	21	535.2	52.8	933	22	ABA68294	Human foetal liver
C	22	535.2	52.8	933	22	AAI23185	Probe #13118 for g
C	23	535.2	52.8	933	22	AAI48508	Probe #17194 used
C	24	535.2	52.8	933	24	ABSI6484	Human genome-deriv
C	25	519.8	51.3	813	24	ABK65147	DNA encoding human
C	26	518.4	51.1	1792	22	AAF58609	DNA encoding human
C	27	517.8	51.1	993	22	AAH32365	Human RECAP polyu
C	28	511.4	50.4	1782	22	AAH32365	Human olfactory re
C	29	481.2	47.5	1713	17	AAT33900	Human cDNA encodin
C	30	481.2	47.5	1713	21	AAA70768	Human G-protein co
C	31	481.2	47.5	1713	21	AAZ60116	Human G-protein co
C	32	429.6	42.4	576	22	AAH32210	Human olfactory re
C	33	415	40.9	952	24	ABQ88400	Human G protein co
C	34	415	40.9	963	24	ABQ88399	Human G protein co
C	35	414	40.8	933	22	AAS42451	Human cDNA encodin
C	36	414	40.8	933	24	ABK37737	DNA encoding human
C	37	409.2	40.4	942	22	AAH32170	DNA encoding human
C	38	407.2	40.2	937	22	AAH32170	Human olfactory re
C	39	405.6	40.0	1411	24	ABK65148	Human olfactory re
C	40	402.2	39.7	930	22	AAS42444	DNA encoding human
C	41	402.2	39.7	930	24	ABK37730	Human cDNA encodin
C	42	398.2	39.3	988	22	AAH32061	Human olfactory re
C	43	390	38.5	933	22	AAS42450	Human cDNA encodin
C	44	390	38.5	933	24	ABK37736	DNA encoding G-cou
C	45	389.8	38.4	938	24	ABN89129	Human GPCR9 nucleo

#### ALIGNMENTS

RESULT 1	AAS09956	standard; DNA: 1014 BP.
ID	AAS09956	
XX	AAS09956;	
AC	24-OCT-2001	(first entry)
XX	DNA encoding human odorant receptor (OR)-like protein, NOV12.	
DE	Odorant receptor; human; OR; NOV12; therapeutic; neuro-olfactory system;	
XX	trauma; surgery; neoplastic disorder; gene therapy; adenocarcinoma;	
KW	lymphoma; prostate cancer; uterus cancer; immune response; asthma; AIDS;	
KW	acquired immunodeficiency syndrome; Crohn's disease; multiple sclerosis;	
KW	Albright hereditary osteodystrophy; diagnostic; ds.	
XX	Homo sapiens.	
OS		
XX		
FH	Key	Location/Qualifiers
FT	5'UTR	1..54
FT		/*tag= a
FT		55..987
FT	CDS	/*tag= b
FT		/product= "Odorant receptor-like protein, NOV12"
FT	3'UTR	988..1014
FT		/*tag= c
PN	WO200151632-A2.	
XX	19-JUL-2001.	
PD		
XX	16-JAN-2001; 2001WO-0501513.	
PF		

XX 13-JAN-2000; 2000US-0175989.  
 PR 14-JAN-2000; 2000US-0176134.  
 PR 25-JAN-2000; 2000US-0177839.  
 PR 26-JAN-2000; 2000US-0178191.  
 PR 26-JAN-2000; 2000US-0178227.  
 PR 14-JUL-2000; 2000US-0218324.  
 PR 24-JUL-2000; 2000US-0220253.  
 PR 25-JUL-2000; 2000US-0220590.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Padigaru M, Prayaga SK, Taupier RJ, Mishra V, Tchernev VT;  
 PI Spytek KA, Li L;  
 XX WPI: 2001-451859/48.  
 DR P-PSDB; AAU05142.  
 DR  
 XX  
 PT New NOVX polypeptides and polynucleotides, useful for treating or  
 PT preventing disorders of the neuro-olfactory system, cancer and multiple  
 PT sclerosis -  
 PS  
 PS Claim 9; Page 55; 141pp; English.  
 XX  
 CC The sequence represents the coding sequence of human odorant receptor  
 CC (OR)-like protein, NOV12. The NOV12 polypeptide, nucleic acid and  
 CC antibody are useful as therapeutics, particularly in the manufacture of a  
 CC medicament for treating a syndrome associated with a human disease, which  
 CC includes a pathology associated with NOV12 polypeptide. The NOV12 nucleic  
 CC acid and polypeptide are especially useful in therapeutic or  
 CC prophylactic applications for disorders of the neuro-olfactory system,  
 CC e.g. those induced by trauma, surgery and/or neoplastic disorders. The  
 CC DNA encoding the protein is useful in gene therapy for treating the  
 CC above conditions. Furthermore, the nucleic acids and polypeptides are  
 CC useful in treating adenocarcinoma, lymphoma, prostate cancer, uterus  
 CC cancer, immune response, acquired immunodeficiency syndrome (AIDS),  
 CC asthma, Crohn's disease, multiple sclerosis or Albright hereditary  
 CC osteodystrophy. These are also useful in developing powerful assay  
 CC system for functional analysis of various human disorders, as well as  
 CC in diagnostic applications.  
 XX  
 SQ Sequence 1014 BP; 219 A; 281 C; 216 G; 298 T; 0 other;

Query Match 97.7%; Score 991; DB 22; Length 1014;  
 Best Local Similarity 99.8%; Pred. No. 4.3e-287;  
 Matches 1013; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 TAAACACTTCTCTTAACCATGAGCAATTAATTGCTCTCTGTCATAGGATATGGGG 60  
 DB 1 TAAACACTTCTCTTAACCATGAGCAATTAATTGCTCTCTGTCATAGGATATGGGG 60  
 QY 61 GACATATATACATCATCAGAGCTTCTCTACTGGGATTTCCCGTGGCCCAAGATT 120  
 DB 61 GACATATATACATCATCAGAGCTTCTCTACTGGGATTTCCCGTGGCCCAAGATT 120  
 QY 121 CAGATGCTCTCTTTGGGCTTCTCCCTGTTCTACGTTCTTACCCCTGCTGGGGAACGGG 180  
 DB 121 CAGATGCTCTCTTTGGGCTTCTCCCTGTTCTACGTTCTTACCCCTGCTGGGGAACGGG 180  
 QY 181 ACCATATCTGGGGCTCATCTCACTGAGCTCCAGACTGCAGCCGCC-TGTACTTCTTCTC 239  
 DB 181 ACCATATCTGGGGCTCATCTCACTGAGCTCCAGACTGCAGCCGCCCATGTACTTCTTCTC 240  
 QY 240 TCACACCTGGCGGCTGTCGACATGCGCTTACGCTGCAACACGGGTGCCCGGATGCTGGTG 299  
 DB 241 TCACACCTGGCGGCTGTCGACATGCGCTTACGCTGCAACACGGGTGCCCGGATGCTGGTG 300  
 QY 300 AACCTCTGTCATCCAGCCCAAGCCCATCTCTTGGCGGCGCATGATGACACCTTTCTG 359  
 DB 301 AACCTCTGTCATCCAGCCCAAGCCCATCTCTTGGCGGCGCATGATGACACCTTTCTG 360  
 QY 360 TTTTCACCTTTTGTGTCACAGAAATGCTCTCTGCTGCTGATGCTCATGATCTGTAC 419  
 DB 360 TTTTCACCTTTTGTGTCACAGAAATGCTCTCTGCTGCTGATGCTCATGATCTGTAC 419

DB 361 TTTCCACTTTTGTGTCACAGAAATGCTCTCTGCTGCTGATGATCCATGATCTGTAC 420  
 QY 420 GTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCAC 479  
 DB 421 GTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCAC 480  
 QY 480 CTCGGGGTGAATCTCTGAGACCACTGAGAGTCTTTTATCCTTGATTCATCTTGTGTACTT 539  
 DB 481 CTCGGGGTGAATCTCTGAGACCACTGAGAGTCTTTTATCCTTGATTCATCTTGTGTACTT 540  
 QY 540 CTACCTTTACCTCTCTGAGGCCCCAGAAATTTATCACTTTTGTGAATCTTGGC 599  
 DB 541 CTACCTTTACCTCTCTGAGGCCCCAGAAATTTATCAC-TTTTGTGAATCTTGGC 599  
 QY 600 TGTCTCAAACTGCTGTGAGATACCCACATCAATGAGAACATGCTTGGCCGAGC 659  
 DB 600 TGTCTCAAACTGCTGTGAGATACCCACATCAATGAGAACATGCTTGGCCGAGC 659  
 QY 660 AATTTCGCGCTGGTGGGACCCCTGTGCCAATTTGATTTGATATGTGCATCTCTG 719  
 DB 660 AATTTCGCGCTGGTGGGACCCCTGTGCCAATTTGATTTGATATGTGCATCTCTG 719  
 QY 720 TGCATCTCTCAGATTCATCAAGGAGATTCAGAGGAAGCCCTCTGCACCTGCTCTC 779  
 DB 720 TGCATCTCTCAGATTCATCAAGGAGATTCAGAGGAAGCCCTCTGCACCTGCTCTC 779  
 QY 780 CCACCTCTGTGTGATTTGAGACTCTTTATGAGCAGCCATATATGATGTTGAGACCAG 839  
 DB 780 CCACCTCTGTGTGATTTGAGACTCTTTATGAGCAGCCATATATGATGTTGAGACCAG 839  
 QY 840 ATATGGGAACCCCAAGAGCAGAGAAATATCTCTCTGCTGTTTCACAGCCTCTTAATCC 899  
 DB 840 ATATGGGAACCCCAAGAGCAGAGAAATATCTCTCTGCTGTTTCACAGCCTCTTAATCC 899  
 QY 900 CATGCTCAATCCCTTATCTGTAGTCTTAGAAGCTCAGAAAGTGAATACCTTGAAGAG 959  
 DB 900 CATGCTCAATCCCTTATCTGTAGTCTTAGAAGCTCAGAAAGTGAATACCTTGAAGAG 959  
 QY 960 AGTCTGGAGTAGAAGGCTTTATGAAAGATTTATGGCATTTGACTGACAA 1014  
 DB 960 AGTCTGGAGTAGAAGGCTTTATGAAAGATTTATGGCATTTGACTGACAA 1014

RESULT 2  
 AAS09955  
 ID AAS09955 standard; DNA: 1012 BP.  
 XX  
 AC AAS09955;  
 XX  
 DT 24-OCT-2001 (first entry)

DE DNA encoding human odorant receptor (OR)-like protein, NOV11.  
 XX  
 KW Odorant receptor; human; OR; NOV11; therapeutic; neuro-olfactory system;  
 KW trauma; surgery; neoplastic disorder; gene therapy; adenocarcinoma;  
 KW lymphoma; prostate cancer; uterus cancer; immune response; asthma; AIDS;  
 KW acquired immunodeficiency syndrome; Crohn's disease; multiple sclerosis;  
 KW Albritght hereditary osteodystrophy; diagnostic; ds.

OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 5'UTR 1..53  
 FT CDS /tag= a  
 FT /tag= b 54..986  
 FT /product= "Odorant receptor-like protein, NOV11"  
 FT 3'UTR 987..1012  
 FT /tag= c  
 PN WO200151632-A2.  
 XX 19-JUL-2001.



XX 19-JUL-2001.  
PD 16-JAN-2001; 2001MO-US01513.  
XX 13-JAN-2000; 2000US-0175989.  
PR 14-JAN-2000; 2000US-0176134.  
PR 25-JAN-2000; 2000US-0177839.  
PR 26-JAN-2000; 2000US-0178191.  
PR 26-JAN-2000; 2000US-0178227.  
PR 14-JUL-2000; 2000US-0218324.  
PR 24-JUL-2000; 2000US-0220253.  
PR 25-JUL-2000; 2000US-0220590.  
XX (CURA-) CURAGEN CORP.  
PI Padigaru M, Prayaga SK, Taupier RJ, Mishra V, Tchernev VT;  
PI Spytek KA, Li L;  
XX WPI; 2001-451859/48.  
DR P-PSDB; AAU05132.  
XX  
XX New NOVX polypeptides and polynucleotides, useful for treating or  
PT preventing disorders of the neuro-olfactory system, cancer and multiple  
PT sclerosis -  
XX  
XX Claim 9; page 15; 141pp; English.  
XX  
XX The sequence represents the coding sequence of human odorant receptor  
CC (OR)-like protein, NOV2. The NOV2 polypeptide, nucleic acid and antibody  
CC are useful as therapeutics, particularly in the manufacture of a  
CC medicament for treating a syndrome associated with a human disease, which  
CC includes a pathology associated with NOV2 polypeptide. The NOVX nucleic  
CC acid and polypeptide are especially useful in therapeutic or  
CC prophylactic applications for disorders of the neuro-olfactory system,  
CC e.g. those induced by trauma, surgery and/or neoplastic disorders. The  
CC DNA encoding the protein is useful in gene therapy for treating the  
CC above conditions. Furthermore, the nucleic acids and polypeptides are  
CC useful in treating adenocarcinoma, lymphoma, prostate cancer, uterus  
CC cancer, immune response, acquired immunodeficiency syndrome (AIDS),  
CC asthma, Crohn's disease, multiple sclerosis or Albright hereditary  
CC osteodystrophy. These are also useful in developing powerful assay  
CC system for functional analysis of various human disorders, as well as  
CC in diagnostic applications.  
XX  
SQ Sequence 1040 BP; 229 A; 285 C; 221 G; 305 T; 0 other;

Query Match 97.0%; Score 983.6; DB 22; Length 1040;  
Best Local Similarity 99.4%; Pred. No. 7.3e-285;  
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

OY 1 TAAACACTTCTCTTAACCATGAGCATTAATGATTTCTCTGTCATAGGATATGGGG 60  
DB 28 TAAACACTTCTCTTAACCATGAGCATTAATGATTTCTCTGTCATAGGATATGGGA 87  
OY 61 GACAATATTAACATCCACAGAGTTCCTCTACTGGATTTCCCGTGGCCCAAGGATT 120  
DB 88 GACAATATTAACATCCACAGAGTTCCTCTACTGGATTTCCCGTGGCCCAAGGATT 147  
OY 121 CAGATGCTCTCTTTGGGCTTCTCTCCCTGTCTACGCTTCAACCTGCTGGGGAACGGG 180  
DB 148 CAGATGCTCTCTTTGGGCTTCTCTCCCTGTCTACGCTTCAACCTGCTGGGGAACGGG 207  
OY 181 ACCATCTGGGGCTCATCTCACTGGACTGCAGACTGCACGCCCC-TGTACTTCTTCTC 239  
DB 208 ACCATCTGGGGCTCATCTCACTGGACTGCAGACTGCACGCCCCCATGTACTTCTTCTC 267  
OY 240 TCACACCTGGGGCTGTCGACATGCGCTACGCGCTGCACACAGGTCGCCGATGCTGTG 299  
DB 268 TCACACCTGGGGCTGTCGACATGCGCTACGCGCTGCACACAGGTCGCCGATGCTGTG 327  
OY 300 AACCTCCTGCATCCAGCCAGCCCATCTCTTTGGCGGCCGATGATGACAGCCTTCTG 359

DB 328 AACCTCCTGCATCCAGCCAGCCCATCTCTTTGGCGGCCGCGCATGATGACAGCCTTCTG 387  
OY 360 TTTTCCACTTTTGGCTGTACAGAAATGTCTCTCTCTGCTGTGTATGCTCTATGATCTGTAC 419  
DB 388 TTTTCCACTTTTGGCTGTACAGAAATGTCTCTCTCTGCTGTGTATGCTCTATGATCTGTAC 447  
OY 420 GTGGCACTGTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCAC 479  
DB 448 GTGGCACTGTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGCATCAC 507  
OY 480 CTCGCGGTGACTTCTCTGACCACTGAGTCTTTATCTTATCTTGAATCATCTGTGTACTT 539  
DB 508 CTCGCGGTGACTTCTCTGACCACTGAGTCTTTATCTTATCTTGAATCATCTGTGTACTT 567  
OY 540 CTACCTTACCCCTCTGTAGGCCCCAGAAAATTTATCACTTTTGTGTGAATCTTGGC 599  
DB 568 CTACCTTACCCCTCTGTAGGCCCCAGAAAATTTATCAC-TTTTGTGTGAATCTTGGC 626  
OY 600 TGTTCGAACCTTGCCCTGTGAGATACCCACATCAATGAGAACATGCTTGGCCGAGC 659  
DB 627 TGTTCGAACCTTGCCCTGTGAGATACCCACATCAATGAGAACATGCTTGGCCGAGC 686  
OY 660 AATTCTGGGCTGTGGGACCTTGTCCACAATTTAGTTTCATATATGTGATCTCTG 719  
DB 687 AATTCTGGGCTGTGGGACCTTGTCCACAATTTAGTTTCATATATGTGATCTCTG 746  
OY 720 TGTATCTCTGAGATCCAAATCAAGGAAAGTTCAAGGAAAGCCTTGTGACCTGCTTC 779  
DB 747 TGTATCTCTGAGATCCAAATCAAGGAAAGTTCAAGGAAAGCCTTGTGACCTGCTTC 806  
OY 780 CCACCTGTGTGATGTGACCTTTTATGGCACACCATTAATCATATGTTGGACCCAG 839  
DB 807 CCACCTGTGTGATGTGACCTTTTATGGCACACCATTAATCATATGTTGGACCCAG 866  
OY 840 ATATGGAAACCCCAAGAGCAGAAATATCTCTGCTGTTTACAGACCTCTTAAATC 899  
DB 867 ATATGGAAACCCCAAGAGCAGAAATATCTCTGCTGTTTACAGACCTCTTAAATC 926  
OY 900 CATGCTCAATCCCTTATCTGTAGCTTTAGGAACCTCAAGAGTGAAGATCTTGAAGAG 959  
DB 927 CATGCTCAATCCCTTATCTGTAGCTTTAGGAACCTCAAGAGTGAAGATCTTGAAGAG 986  
OY 960 AGTGTGGAGTAGAAAGGCTTTATGAAAGAGATTTGACATTTGTGACTGACA 1013  
DB 987 AGTGTGGAGTAGAAAGGCTTTATGAAAGAGATTTGACATTTGTGACTGACA 1040

RESULT 4  
ABA09073/c  
ID ABA09073 standard; cDNA; 1315 BP.

XX ABA09073;  
XX AC  
XX 11-JAN-2002 (first entry)  
XX Human olfactory receptor homologue-encoding cDNA, SEQ ID NO:849.  
DE Human: cytokine; cell proliferation; cell differentiation; growth factor;  
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenes;  
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; anti-inflammatory;  
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
KW cytotatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
XX antifungal; vulnery; antitumor; ss.  
OS Homo sapiens.  
XX



PN WO200157188-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001MO-US03800.  
XX  
PR 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-0560875.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-457740/49.  
DR P-PSDB; ABB11829.  
XX  
PT Human proteins and DNA encoding sequences useful for preventing,  
PT treating or ameliorating a medical condition in a mammalian subject  
PT e.g. arthritis and cancer -  
XX  
PS Claim 1; Page 753; 1963pp; English.  
XX  
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, thereby  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis, cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a cDNA encoding a  
CC novel human polypeptide of the invention.  
XX  
SQ Sequence 1315 BP; 380 A; 281 C; 360 G; 294 T; 0 other;

Query Match 92.9%; Score 942; DB 22; Length 1315;  
Best Local Similarity 99.6%; Pred. No. 2.6e-272;  
Matches 986; Conservative 0; Mismatches 0; Indels 4; Gaps 4;

QY 1 TAAACACTTCTCCTAACCAGATGAGCATTAATTGATTCCTCTGCATAGGATATGGG 60  
|||||  
Db 989 TAAACACTTCTCCTAACCAGATGAGCATTAATTGATTCCTCTGCATAGGATATGGG 930  
|||||

QY 61 GACATATATACATGCATCAGAGTTCCTCCTACTGGGATTTCCGTTGGCCCAAGATT 120  
|||||  
Db 929 GACATATATACATGCATCAGAGTTCCTCCTACTGGGATTTCCGTTGGCCCAAGATT 870  
|||||

QY 121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCACGCTTCACCCCTGGGGAACGG 180  
|||||  
Db 869 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTCACGCTTCACCCCTGGGGAACGG 810  
|||||

QY 181 ACCATAGTGGGCTCATCTACATGAGACTCAGACTGACGCCCCC-TGTACTTCTTCTC 239  
|||||  
Db 809 ACCATAGTGGGCTCATCTACATGAGACTCAGACTGACGCCCCCATGACTTCTTCTC 750  
|||||

QY 240 TCACACCT-GGCGTCTGACATGCGCTACGCGCTGCAACACGGTGCCCGGATGCTGGT 298  
|||||  
Db 749 TCACACCTGGCGGCTGCTGACATGCGCTACGCGCTGCAACACGGTGCCCGGATGCTGGT 690  
|||||

QY 299 GAACCTCTGCATCCAGCCCAAGCCCATCTCTTGGGCGCCGATGATGACACCTTCT 358  
|||||  
Db 689 GAACCTCTGCATCCAGCCCAAGCCCATCTCTTGGGCGCCGATGATGACACCTTCT 630  
|||||

QY 359 GTTTCACATTTTGTCTGTACAGAAATGCTCTCTCTGCTGCTGATGATGATGATGAT 418  
|||||  
Db 629 GTTTCACATTTTGTCTGTACAGAAATGCTCTCTCTGCTGCTGATGATGATGATGAT 570  
|||||

QY 419 GGT-GGACATCTGACACCCCTCCGATATTTGGCCATGATGATGATGATGATGATGAT 477  
|||||  
Db 569 GGTGGGCAATCTGACACCCCTCCGATATTTGGCCATGATGATGATGATGATGATGAT 510  
|||||

QY 478 CCTGGGCTGACTTCTCTGACACATGAGAGTCTTTTATCTTATCTTATCTTATCTTATCT 537  
|||||  
Db 509 CCTGGGCTGACTTCTCTGACACATGAGAGTCTTTTATCTTATCTTATCTTATCTTATCT 450  
|||||

QY 538 TTCTACCTTACCTTCTCTGAGGCCCCAGAAATTTATCTTATCTTATCTTATCTTATCT 597  
|||||  
Db 449 TTCTACCTTACCTTCTCTGAGGCCCCAGAAATTTATCTTATCTTATCTTATCTTATCT 391  
|||||

QY 598 GCTGTCTCAAACTTGCTGTGACATACCCACATCAAGAAACATGATGATGATGATGAT 657  
|||||  
Db 390 GCTGTCTCAAACTTGCTGTGACATACCCACATCAAGAAACATGATGATGATGATGAT 331  
|||||

QY 658 GCAATTTCTGGGCTGTGGAGCCCTGTCCACATTTGATTTATATATGATGATGATGAT 717  
|||||  
Db 330 GCAATTTCTGGGCTGTGGAGCCCTGTCCACATTTGATTTATATATGATGATGATGAT 271  
|||||

QY 718 TGTGCTATCTCAGATCCAAATCAAGGAAAGTTCAGAGGAAAGCTTCTGACCTGCTTC 777  
|||||  
Db 270 TGTGCTATCTCAGATCCAAATCAAGGAAAGTTCAGAGGAAAGCTTCTGACCTGCTTC 211  
|||||

QY 778 TCCACCTCTGTGTGATTTGACCTTTTATGACACAGCCATTTATGATGATGATGATGAT 837  
|||||  
Db 210 TCCACCTCTGTGTGATTTGACCTTTTATGACACAGCCATTTATGATGATGATGATGAT 151  
|||||

QY 838 AGATATGGGAACCCCAAGGAGCAGAAATATCTCTGCTGTTTTCACAGCCCTTTAT 897  
|||||  
Db 150 AGATATGGGAACCCCAAGGAGCAGAAATATCTCTGCTGTTTTCACAGCCCTTTAT 91  
|||||

QY 898 CCCATGCTCAATCCCTTATCTGATGCTTTAGGAACCTCAAGAGTGAAGATATCTTGAAG 957  
|||||  
Db 90 CCCATGCTCAATCCCTTATCTGATGCTTTAGGAACCTCAAGAGTGAAGATATCTTGAAG 31  
|||||

QY 958 AGAGTGTGGAGTAGAAAGGCTTTATGA 987  
|||||  
Db 30 AGAGTGTGGAGTAGAAAGGCTTTATGA 1

RESULT 5  
AAF58613  
ID AAF58613 standard; cDNA; 974 BP.  
XX  
AC AAF58613;  
XX  
AC  
XX  
DT 24-APR-2001 (first entry)  
XX  
DE Human RECAP polynucleotide, SEQ ID NO: 41.  
XX  
KW Human; RECAP; receptors and associated proteins; cerebroprotective;

KW nootropic; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;  
KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;  
KW antithyroid; immunosuppressive; nephrotropic; antileprosy; thymomimetic;  
KW cytostatic; antibacterial; virucide; fungicide; protozoacide;  
KW antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer; ss.  
XX Homo sapiens.  
OS  
XX  
PN WO200107612-A2.  
XX  
PD 01-FEB-2001.  
XX  
PF 21-JUL-2000; 2000WO-US20035.  
XX  
PR 21-JUL-1999; 99US-0145232.  
PR 07-OCT-1999; 99US-0158578.  
PR 12-NOV-1999; 99US-0165192.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Au-Young J, Bandman O, Tang YT, Yue H, Azimzal Y, Burford N;  
PI Baughn MR, Lu DAM, Hillman JL, Patterson C, Lal P;  
XX  
DR WPI; 2001-168554/17.  
DR P-PSDB; AAB68889.  
XX  
PT Novel receptors and associated proteins for diagnosis and treatment of  
PT neurological disorders, immunological disorders including autoimmune/  
PT inflammatory disorders and cell proliferative disorders such as cancer  
PT  
XX  
PS Claim 5; Page 125-126; 128pp; English.  
XX  
CC The present sequence encodes a human RECAP (receptors and associated  
CC proteins) polypeptide. RECAP polynucleotides and polypeptides are useful  
CC in the diagnosis, treatment and prevention of neurological disorders  
CC such as stroke, Alzheimer's disease, Pick's disease, Huntington's  
CC disease, dementia, Parkinson's disease, Down's syndrome, amyotrophic  
CC lateral sclerosis, multiple sclerosis, bacterial and viral meningitis,  
CC CJD (Creutzfeldt-Jakob disease), GSS (Gerstmann -Strausler-Scheinker  
CC syndrome); immunological disorders, including autoimmune/inflammatory  
CC disorders such as AIDS, Digeorge's syndrome, severe combined  
CC immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's  
CC disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,  
CC diabetes mellitus, Good pasture's syndrome, gout, Grave's disease,  
CC Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,  
CC bacterial, fungal, parasitic, protozoal, and helminthic infections; and  
CC cell proliferation disorders such as arteriosclerosis, atherosclerosis,  
CC cirrhosis, hepatitis and cancer.  
XX  
SQ Sequence 974 BP; 206 A; 270 C; 214 G; 284 T; 0 other;  
Query Match 92.3%; Score 936; DB 22; Length 974;  
Best Local Similarity 99.8%; Pred. No. 1.4e-270;  
Matches 958; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
QY 55 ATGGGGACATATATACATCCATCAGAGTTCCTCTACTGGGATTTCCCGTGGCCCA 114  
Db 1 ATGGGGACATATATACATCCATCAGAGTTCCTCTACTGGGATTTCCCGTGGCCCA 60  
QY 115 AGGATTCAGATGCTCTTTGGGCTTCTCTCCCTGTTCAGTCTTCACCCGTGGGG 174  
Db 61 AGGATTCAGATGCTCTTTGGGCTTCTCTCCCTGTTCAGTCTTCACCCGTGGGG 120  
QY 175 AACGGACCATACTGGGCTCATCTCAGACTCAGACTCAGCCGCC-CTGACTTC 233  
Db 121 AACGGACCATACTGGGCTCATCTCAGACTCAGACTCAGCCGCCCATGTACTTC 180  
QY 234 TTCCTCTCAACCTGGCGTCTGAGCATGCGCTAGCGCTGCAACACGGTGCCTGGATG 293  
Db 181 TTCCTCTCAACCTGGCGTCTGAGCATGCGCTAGCGCTGCAACACGGTGCCTGGATG 240  
QY 294 CTGTGAACCTCTGCATCCAGCAAGCCCATCTCTTTGGGGCCGCATGATGACAGCC 353

Db 241 CTGTGAACCTCTGCATCCAGCAAGCCCATCTCTTTGCGGGCCGATGATGACAGCC 300  
QY 354 TTTCGTGTTTCCACTTTTGTGCTGTACAGAAATGCTCCTCTCTGGTGGATGCTCTATGAT 413  
Db 301 TTTCGTGTTTCCACTTTTGTGCTGTACAGAAATGCTCCTCTCTGGTGGATGCTCTATGAT 360  
QY 414 CTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCTGC 473  
Db 361 CTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCTGC 420  
QY 474 ATCACCCCTCGCGGTGACTTCTCTGAGACCACCTGAGTCTTTATCTGATTCATCTTGTG 533  
Db 421 ATCACCCCTCGCGGTGACTTCTCTGAGACCACCTGAGTCTTTATCTGATTCATCTTGTG 480  
QY 534 TTACTTCTACCTTTACCCCTCTCTGTAGCCCCAGAAATTTATCACTTTTGTGAAAT 593  
Db 481 TTACTTCTACCTTTACCCCTCTCTGTAGCCCCAGAAATTTATCAC-TTTTTTGTGAAAT 539  
QY 594 CTGGGCTGTCTCAAACTTGCTGTGAGATACCCATCAATGAGAACATGGTCTTGGC 653  
Db 540 CTGGGCTGTCTCAAACTTGCTGTGAGATACCCATCAATGAGAACATGGTCTTGGC 599  
QY 654 CGGAGCAATTTCTGGGCTGTGAGACCCCTGTGCACAAATGTAGTTTCAATATATGTGCAT 713  
Db 600 CGGAGCAATTTCTGGGCTGTGAGACCCCTGTGCACAAATGTAGTTTCAATATATGTGCAT 659  
QY 714 CCTCTGTCTATCTTCAAGATCCAAATCAAGGGAAGTTCAGAGAAAGCCTTCTGCACCTG 773  
Db 660 CCTCTGTCTATCTTCAAGATCCAAATCAAGGGAAGTTCAGAGAAAGCCTTCTGCACCTG 719  
QY 774 CTTCGCCACCTCTGTGTGATGAGCTCTTTATGGCACGCCATTTATCATGTATGTTGG 833  
Db 720 CTTCGCCACCTCTGTGTGATGAGCTCTTTATGGCACGCCATTTATCATGTATGTTGG 779  
QY 834 ACCCAGATATGGGAACCCCAAGGACAGAAATATCTCTCTGTTTTCACAGCCTCTT 893  
Db 780 ACCCAGATATGGGAACCCCAAGGACAGAAATATCTCTCTGTTTTCACAGCCTCTT 839  
QY 894 TAAATCCATGCTCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAAGTGAAGAAATCTT 953  
Db 840 TAAATCCATGCTCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAAGTGAAGAAATCTT 899  
QY 954 GAAGAGAGTCTGGGAGTAGAAGGCTTATGAAGAAGATTATGATGTGACTGACA 1013  
Db 900 GAAGAGAGTCTGGGAGTAGAAGGCTTATGAAGAAGATTATGATGTGACTGACA 959  
RESULT 6  
AAH32366 standard; DNA; 930 BP.  
ID AAH32366  
XX  
AC AAH32366;  
XX  
DT 30-JUL-2001 (first entry)  
XX  
DE Human olfactory receptor polynucleotide, SEQ ID NO: 939.  
XX  
KW Human; olfactory receptor; OR; primary scent determination;  
KW secondary scent determination; polypeptide library; odour receptor;  
KW scent profile; scent fingerprint; scent representation; ds.  
XX  
OS Homo sapiens.  
OS  
PN WO200127158-A2.  
PN  
PD 19-APR-2001.  
PD  
XX  
PF 06-OCT-2000; 2000WO-US27582.  
PF  
XX  
PR 08-OCT-1999; 99US-0158615.  
PR  
PR 24-FEB-2000; 2000US-0184809.  
XX



PA (DIGI-) DIGISCENTS.  
PA (YEDA ) YEDA RES & DEV CO LTD.

XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
DR WPI; 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists -  
XX Claim 8; Page 554; 1857pp; English.

XX The present sequence is one of a number of isolated polynucleotides  
CC which encode polypeptides involved in olfactory sensation. The  
CC polynucleotides can be used in screening for olfactory agonists and  
CC antagonists. The methods allow for the determination of primary  
CC scents and the identification of the odour receptors used to detect  
CC these primary scents. The methods also enable determination of  
CC secondary scents and the identification of combinations of odour  
CC receptors that are involved in detecting such secondary scents.  
CC This enables the construction of a scent representation (also called  
CC a scent fingerprint or scent profile), which may be used to re-create  
CC and edit scents. Libraries of olfactory receptors are useful for  
CC determining the interaction pattern of a composition with the receptors,  
CC and can be used for determining differences in the olfactory faculties  
CC of different individuals.

XX SQ Sequence 930 BP; 192 A; 266 C; 201 G; 271 T; 0 other;

Query Match 89.3%; Score 905.4; DB 22; Length 930;  
Best Local Similarity 99.7%; Pred. No. 2.1e-261;  
Matches 928; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

OY 55 ATGGGGACAAATATACATCCATGACAGAGTTCTCTACTGGGATTTCCCGTGGCCCA 114  
DB 1 ATGGGGACAAATATACATCCATGACAGAGTTCTCTACTGGGATTTCCCGTGGCCCA 60  
OY 115 AGGATTCAGATGCTCTCTTTGGGCTCTTCCCTGTTTACGTTTCAACCTGGTGGG 174  
DB 61 AGGATTCAGATGCTCTCTTTGGGCTCTTCCCTGTTTACGTTTCAACCTGGTGGG 120  
OY 175 AACGGACCACTACTGGGGCTCATCTGACTGCACTCCAGACTGACGCCCC-TCGACTTC 233  
DB 121 AACGGACCACTACTGGGGCTCATCTGACTGCACTCCAGACTGACGCCCCCATGACTTC 180  
OY 234 TTCCTCTACACCTGGCGCTGCTGACATCGCCTACGCTGCAACACAGGTGCCCGGATG 293  
DB 181 TTCCTCTACACCTGGCGCTGCTGACATCGCCTACGCTGCAACACAGGTGCCCGGATG 240  
OY 294 CTGTGAACCTCTGCTGCTGACAGCCATCTCTTGGGGGCCGATGATGACAGCC 353  
DB 241 CTGTGAACCTCTGCTGCTGACAGCCATCTCTTGGGGGCCGATGATGACAGCC 300  
OY 354 TTTCTGTTTCCACTTTTGTGTCAGAGATGTCCTCTGTTGTTGTTGTTGTTGTTGTT 413  
DB 301 TTTCTGTTTCCACTTTTGTGTCAGAGATGTCCTCTGTTGTTGTTGTTGTTGTTGTT 360  
OY 414 CTGTAGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGC 473  
DB 361 CTGTAGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGC 420  
OY 474 ATCACCTCGGGGTGACTTCTGACACCACTGAGTCTTATCTGATTCATCTTGTG 533  
DB 421 ATCACCTCGGGGTGACTTCTGACACCACTGAGTCTTATCTGATTCATCTTGTG 480  
OY 534 TTAGTCTTACCTTACCTTCTGTAGGCCCCAGAAAATTTATCACTTTTGTGCAAT 593  
DB 481 TTAGTCTTACCTTACCTTCTGTAGGCCCCAGAAAATTTATCAC-TTTTGTGCAAT 539  
OY 594 CTGCTGTCTCAAACTTGCCCTGTCAGAGATACCCACATCAATGAGACATGCTTGGC 653  
DB 540 CTGCTGTCTCAAACTTGCCCTGTCAGAGATACCCACATCAATGAGACATGCTTGGC 599

OY 654 CGGAGCAATTTCTGGGCTGTGGGACCCCTGTGTCACAATGTAGTTGATATATGTGCAT 713  
DB 600 CGGAGCAATTTCTGGGCTGTGGGACCCCTGTGTCACAATGTAGTTGATATATGTGCAT 659  
OY 714 CCTCTGTCTATCTTCCATTCAGATCCAAATCAAGGGAAGTTGAGGAAAGCTTCTGCACCTG 773  
DB 660 CCTCTGTCTATCTTCCATTCAGATCCAAATCAAGGGAAGTTGAGGAAAGCTTCTGCACCTG 719  
OY 774 CTCTCCACCTCTGTGTGATGAGTGGACTCTTTATGACACAGCCATTCATGATGTTGG 833  
DB 720 CTCTCCACCTCTGTGTGATGAGTGGACTCTTTATGACACAGCCATTCATGATGTTGG 779  
OY 834 ACCCATATATGGGAAACCCAGAGAGCAGAGAAATATCTCTGCTGTTTCACAGCCTCTT 893  
DB 780 ACCCATATATGGGAAACCCAGAGAGCAGAGAAATATCTCTGCTGTTTCACAGCCTCTT 839  
OY 894 TAATCCATGCTCAATCCCTTATCTGTAGTCTTAGGAATCAGAAAGTGAAGATACTTT 953  
DB 840 TAATCCATGCTCAATCCCTTATCTGTAGTCTTAGGAATCAGAAAGTGAAGATACTTT 899  
OY 954 GAAGAGTGTCTGGAGTAGAAGGCTTTA 984  
DB 900 GAAGAGTGTCTGGAGTAGAAGGCTTTA 930

## RESULT 7

ABK65141  
ID ABK65141 standard; cDNA; 933 BP.

AC ABK65141;  
XX  
DT 02-JUL-2002 (first entry)

XX cDNA encoding human PHOR1-F5D6.

XX Human; PHOR1-A11; PHOR1-F5D6; prostate cancer: cytostatic; gene; ss.

XX Homo sapiens.

XX W0200214501-A2.

XX 21-FEB-2002.

XX 17-AUG-2001; 2001WO-US25862.

XX 17-AUG-2000; 2000US-226241P.

XX (AGEN-) AGENSYS INC.

XX Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;  
PI Jakobovits A;

XX WPI; 2002-269193/31.

XX P-PSDB; AAU91543.

XX Monitoring PHOR1-A11/PHOR1-F5D6 gene products for monitoring presence  
PT of cancer in subject, by determining status of PHOR1-A11/PHOR1-F5D6  
PT gene products in tissue sample from subject and comparing it to normal  
PT sample

XX Claim 43; Fig 2b; 250pp; English.

XX The present invention relates to the isolation of novel human genes  
CC designated PHOR1-A11 and PHOR1-F5D6 and their encoded proteins. The  
CC gene encoding PHOR1-A11 maps to chromosome 1q23, and the gene encoding  
CC PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-A11 and PHOR1-F5D6  
CC polynucleotide and polypeptide sequences are useful in diagnostic and  
CC therapeutic methods, and compositions for various cancers such as  
CC prostate cancer. The sequences are useful for inhibiting the growth of  
CC cancer cells that express PHOR1-A11 or PHOR1-F5D6 and for treating  
CC cancer. The PHOR1-A11 or PHOR1-F5D6 polypeptide or a fragment thereof  
CC can be used to elicit an immune response. The present sequence encodes  
CC human PHOR1-F5D6.

XX Sequence 933 BP; 194 A; 265 C; 203 G; 271 T; 0 other;  
 SQ

Query Match 89.1%; Score 903.6; DB 24; Length 933;  
 Best Local Similarity 99.4%; Pred. No. 7.4e-261;  
 Matches 928; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 55 ATGGGGGACATATATACATCCATCAGAGAGTTCCTCTACTGGGATTTCCCGTGGCCCA 114  
 DB 1 ATGGGAGACATATATACATCCATCAGAGAGTTCCTCTACTGGGATTTCCCGTGGCCCA 60  
 QY 115 AGGATTCAGATGCTCCTCTTTGGGCTTTCTCCCTGTCTACGCTCTTCACCCCTGCTGGG 174  
 DB 61 AGGATTCAGATGCTCCTCTTTGGGCTTTCTCCCTGTCTACGCTCTTCACCCCTGCTGGG 120  
 QY 175 AACGGGACCATCTGGGGCTCATCTCAGCTGAGCTCCAGACTGACGCCCC--TGTACTTC 233  
 DB 121 AACGGGACCATCTGGGGCTCATCTCAGCTGAGCTCCAGACTGACGCCCCCATGTACTTC 180  
 QY 234 TTCCCTCTCAGACCTGGGGCTGCTGACATCGCCTACGCCCTGCAACACGAGTGCCCGGATG 293  
 DB 181 TTCCCTCTCAGACCTGGGGCTGCTGACATCGCCTACGCCCTGCAACACGAGTGCCCGGATG 240  
 QY 294 CTGGTGAACCTCTCTGCATCCAGCCAGCCCATCTCTTGGGGCCGCGCATGATGACAGAC 353  
 DB 241 CTGGTGAACCTCTCTGCATCCAGCCAGCCCATCTCTTGGGGCCGCGCATGATGACAGAC 300  
 QY 354 TTCTGTCTTCCACTTTTGGCTGTACAGAAATGTCTCCTCTGCTGGTGATGTCTATGAT 413  
 DB 301 TTCTGTCTTCCACTTTTGGCTGTACAGAAATGTCTCCTCTGCTGGTGATGTCTATGAT 360  
 QY 414 CTGTACGTGGCCATCTGACACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGC 473  
 DB 361 CTGTACGTGGCCATCTGACACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGC 420  
 QY 474 ATCACCCCTGCGGGTGAATCTCTGAGACCACTGAGTCTCTTATCTGTATTCATCTCTGTG 533  
 DB 421 ATCACCCCTGCGGGTGAATCTCTGAGACCACTGAGTCTCTTATCTGTATTCATCTCTGTG 480  
 QY 534 TTACTTCTACCTTTACCCCTCTCTGAGAGCCCGCAGAAATTTATCAGCTTTTGTGTAAT 593  
 DB 481 TTACTTCTACCTTTACCCCTCTCTGAGAGCCCGCAGAAATTTATCAGCTTTTGTGTAAT 539  
 QY 594 CTGGCTGTCTCAACACTGCGCTGAGATACCCACATCAATGAGACATGCTCTGGC 653  
 DB 540 CTGGCTGTCTCAACACTGCGCTGAGATACCCACATCAATGAGACATGCTCTGGC 599  
 QY 654 CGGAGCAATTTCTGGGCTGTGGGACCCCTTGTCCACAATTTAGTTTCATATATGTGCAT 713  
 DB 600 CGGAGCAATTTCTGGGCTGTGGGACCCCTTGTCCACAATTTAGTTTCATATATGTGCAT 659  
 QY 714 CCTGTGTCTATCTCTCAGATCCATCAAGGGAAGTTCAGAGGAAAGCCTCTGCACCTG 773  
 DB 660 CCTGTGTCTATCTCTCAGATCCATCAAGGGAAGTTCAGAGGAAAGCCTCTGCACCTG 719  
 QY 774 CTCTCCCACTCTGTGTGATGGAATCTTTATGGCACAGCCATTTATCATGTATGTTGG 833  
 DB 720 CTCTCCCACTCTGTGTGATGGAATCTTTATGGCACAGCCATTTATCATGTATGTTGG 779  
 QY 834 ACCAGATATGGGAACCCCAAGGACGAGAAGAATATCTCTGCTGTTACAGCCCTCTT 893  
 DB 780 ACCAGATATGGGAACCCCAAGGACGAGAAGAATATCTCTGCTGTTACAGCCCTCTT 839  
 QY 894 TAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAAGTTCAGAGTGAAGATATCTTT 953  
 DB 840 TAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAAGTTCAGAGTGAAGATATCTTT 899  
 QY 954 GAAGAGAGTGTCTGGAGTAGAAGGCGCTTATGA 987  
 DB 900 GAAGAGAGTGTCTGGAGTAGAAGGCGCTTATGA 933

RESULT 8

AAH31617  
 ID AAH31617 standard; DNA; 930 BP.

XX AC AAH31617;  
 XX DT 30-JUL-2001 (first entry)  
 XX DE Human olfactory receptor polynucleotide, SEQ ID NO: 190.

KW Human; olfactory receptor; OR; primary scent determination;  
 KW secondary scent determination; polypeptide library; odour receptor;  
 KW scent profile; scent fingerprint; scent representation; ds.  
 OS Homo sapiens.  
 XX WO200127158-A2.  
 XX PD 19-APR-2001.  
 XX PF 06-OCT-2000; 2000WO-US27582.  
 XX PR 08-OCT-1999; 99US-0158615.  
 XX PR 24-FEB-2000; 2000US-0184809.  
 PA (DIGI-) DIGISCENTS.  
 PA (YEDA) YEDA RES & DEV CO LTD.  
 XX PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
 DR WPI; 2001-290713/30.  
 XX PT New polynucleotides which encode polypeptides involved in olfactory  
 PT sensation for identifying olfactory agonists and antagonists -  
 PS Claim 8; page 251; 1857pp; English.

CC The present sequence is one of a number of isolated polynucleotides  
 CC which encode polypeptides involved in olfactory sensation. The  
 CC polynucleotides can be used in screening for olfactory agonists and  
 CC antagonists. The methods allow for the determination of primary  
 CC scents and the identification of the odour receptors used to detect  
 CC these primary scents. The methods also enable determination of  
 CC secondary scents and the identification of combinations of odour  
 CC receptors that are involved in detecting such secondary scents.  
 CC This enables the construction of a scent representation (also called  
 CC a scent fingerprint or scent profile), which may be used to re-create  
 CC and edit scents. Libraries of olfactory receptors are useful for  
 CC determining the interaction pattern of a composition with the receptors,  
 CC and can be used for determining differences in the olfactory faculties  
 CC of different individuals.

XX SQ Sequence 930 BP; 193 A; 265 C; 202 G; 270 T; 0 other;

Query Match 88.8%; Score 900.6; DB 22; Length 930;  
 Best Local Similarity 99.4%; Pred. No. 5.9e-260;  
 Matches 925; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 55 ATGGGGGACATATATACATCCATCAGAGAGTTCCTCTACTGGGATTTCCCGTGGCCCA 114  
 DB 1 ATGGGAGACATATATACATCCATCAGAGAGTTCCTCTACTGGGATTTCCCGTGGCCCA 60  
 QY 115 AGGATTCAGATGCTCCTCTTTGGGCTTTCTCCCTGTCTACGCTTCACCCCTGCTGGG 174  
 DB 61 AGGATTCAGATGCTCCTCTTTGGGCTTTCTCCCTGTCTACGCTTCACCCCTGCTGGG 120  
 QY 175 AACGGGACCATCTAGGGCTCATCTCAGTGGACTCCAGACTGACAGGCCCC--TGTACTTC 233  
 DB 121 AACGGGACCATCTAGGGCTCATCTCAGTGGACTCCAGACTGACAGGCCCCCATGTACTTC 180  
 QY 234 TTCCCTCTCAGACCTGGGGCTGTGACATCGCCTACGCGCTGCAACAGGCTGCCCGGATG 293  
 DB 181 TTCCCTCTCAGACCTGGGGCTGTGACATCGCCTACGCGCTGCAACAGGCTGCCCGGATG 240

OY 294 CTGTGAACCTCTGCATCCAGCCAGCCCATCTCTTGGGGCCGATGATGACAGACC 353  
|||  
DB 241 CTGTGAACCTCTGCATCCAGCCAGCCCATCTCTTGGGGCCGATGATGACAGACC 300  
OY 354 TTCTGTTCACACTTTGCTGTGCAGAAATGTCCTCTGGTGTGATGTCCTATGAT 413  
|||  
DB 301 TTCTGTTCACACTTTGCTGTGCAGAAATGTCCTCTGGTGTGATGTCCTATGAT 360  
OY 414 CTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGCTGC 473  
|||  
DB 361 CTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGCTGC 420  
OY 474 ATGACCCCTGGGCTGACTTCTTGAGACCACTGGAGTCTTTATCCTTGATTCATCTGTG 533  
|||  
DB 421 ATGACCCCTGGGCTGACTTCTTGAGACCACTGGAGTCTTTATCCTTGATTCATCTGTG 480  
OY 534 TTACTTTACCTTTACCTTTCTGTAGGCCCCAGAAAATTTATCAGCTTTTGTGAAAT 593  
|||  
DB 481 TTACTTTACCTTTACCTTTCTGTAGGCCCCAGAAAATTTATCAGCTTTTGTGAAAT 539  
OY 594 CTGGGCTGTCTCAAACTTGCCCTGTGCAGATACCCATCATGAGAAACATGGTCTTGGC 653  
|||  
DB 540 CTGGGCTGTCTCAAACTTGCCCTGTGCAGATACCCATCATGAGAAACATGGTCTTGGC 599  
OY 654 CGGAGCAATTTCTGGGCTGGTGGGACCCCTGTGCACCAATTTAGTTTATATATGTGCAT 713  
|||  
DB 600 CGGAGCAATTTCTGGGCTGGTGGGACCCCTGTGCACCAATTTAGTTTATATATGTGCAT 659  
OY 714 CCTGTGTCTATCTCTGCATCCATCAAGGGAAGTTCAGAGGAAGCCTTCTGCACCTG 773  
|||  
DB 660 CCTGTGTCTATCTCTGCATCCATCAAGGGAAGTTCAGAGGAAGCCTTCTGCACCTG 719  
OY 774 CTTCCTCCACCTCTGTGTGATTTGAGCTCTTTATGACACAGCCATTCATATGTTGG 833  
|||  
DB 720 CTTCCTCCACCTCTGTGTGATTTGAGCTCTTTATGACACAGCCATTCATATGTTGG 779  
OY 834 ACCGAGATATGGGAACCCCAAGGAGCAGAGAATAATCTCTGCTGTTCACAGCCTCTT 893  
|||  
DB 780 ACCGAGATATGGGAACCCCAAGGAGCAGAGAATAATCTCTGCTGTTCACAGCCTCTT 839  
OY 894 TAATCCATGCTCAATCCCTTATCTGTAGTCTTGAAGAACTCAGAAAGTGAATACTTT 953  
|||  
DB 840 TAATCCATGCTCAATCCCTTATCTGTAGTCTTGAAGAACTCAGAAAGTGAATACTTT 899  
OY 954 GAAGAGAGTGTGGAGTAGAAAGGCTTTA 984  
|||  
DB 900 GAAGAGAGTGTGGAGTAGAAAGGCTTTA 930  
RESULT 9  
AAH31648  
ID AAH31648 standard; DNA; 930 BP.  
XX AC AAH31648;  
XX 30-JUL-2001 (first entry)  
DE Human olfactory receptor polynucleotide, SEQ ID NO: 221.  
XX  
KW Human; olfactory receptor; OR; primary scent determination;  
KW secondary scent determination; polypeptide library; odour receptor;  
KW scent profile; scent fingerprint; scent representation; ds.  
XX OS Homo sapiens.  
XX PN WO200127158-A2.  
XX PD 19-APR-2001.  
XX PF 06-OCT-2000; 2000WO-US27582.  
XX PR 08-OCT-1999; 99US-0158615.  
PR 24-FEB-2000; 2000US-0184809.

XX (DIGI-) DIGISCENTS.  
PA (YEDA ) YEDA RES & DEV CO LTD.  
XX  
PI Bellenson J, Smith D, Lancel D, Glusman G, Fuchs T, Yanai I;  
XX WPI: 2001-290713/30.  
DR  
XX  
PT New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists -  
XX  
PS Claim 8; Page 264-265; 1857pp; English.  
XX  
CC The present sequence is one of a number of isolated polynucleotides  
CC which encode polypeptides involved in olfactory sensation. The  
CC polynucleotides can be used in screening for olfactory agonists and  
CC antagonists. The methods allow for the determination of primary  
CC scents and the identification of the odour receptors used to detect  
CC these primary scents. The methods also enable determination of  
CC secondary scents and the identification of combinations of odour  
CC receptors that are involved in detecting such secondary scents.  
CC This enables the construction of a scent representation (also called  
CC a scent fingerprint or scent profile), which may be used to re-create  
CC and edit scents. Libraries of olfactory receptors are useful for  
CC determining the interaction pattern of a composition with the receptors,  
CC and can be used for determining differences in the olfactory faculties  
CC of different individuals.  
XX  
SQ Sequence 930 BP; 193 A; 265 C; 202 G; 270 T; 0 other;  
Query Match 88.8%; Score 900.6; DB 22; Length 930;  
Best Local Similarity 99.4%; Pred. No. 5.9e-260;  
Matches 925; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
OY 55 ATGGGGACAATATACATTCATCAGAGAGTTCCTCTACTGGGATTTCCCGTGGCCA 114  
|||  
DB 1 ATGGGAGACAATATACATTCATCAGAGAGTTCCTCTACTGGGATTTCCCGTGGCCA 60  
OY 115 AGGATTCAGATGCTCTCTTTGGGCTCTCTCCCTGTCTACGTCCTCACCCTGTGGG 174  
|||  
DB 61 AGGATTCAGATGCTCTCTTTGGGCTCTCTCCCTGTCTACGTCCTCACCCTGTGGG 120  
OY 175 AACGGACCATACATGGGGCTCATCTCAGTGGACTCCAGACTGCAGCCGCC -TGTACTTC 233  
|||  
DB 121 AACGGACCATACATGGGGCTCATCTCAGTGGACTCCAGACTGCAGCCGCCCATGTACTTC 180  
OY 234 TTCCCTCTACACCTGGCGGCTGCTGCACATGCCCTACGCCCAACAGGCTGCCCGGATG 293  
|||  
DB 181 TTCCCTCTACACCTGGCGGCTGCTGCACATGCCCTACGCCCAACAGGCTGCCCGGATG 240  
OY 294 CTGTGAACCTCTGCATCCAGCCAGCCCATCTCTTGGGGCCGATGATGACAGACC 353  
|||  
DB 241 CTGTGAACCTCTGCATCCAGCCAGCCCATCTCTTGGGGCCGATGATGACAGACC 300  
OY 354 TTCTGTTCACACTTTGCTGTGCAGAAATGTCCTCTGGTGTGATGTCCTATGAT 413  
|||  
DB 301 TTCTGTTCACACTTTGCTGTGCAGAAATGTCCTCTGGTGTGATGTCCTATGAT 360  
OY 414 CTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGCTGC 473  
|||  
DB 361 CTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGCTGC 420  
OY 474 ATGACCCCTGGGCTGACTTCTTGAGACCACTGGAGTCTTTATCCTTGATTCATCTGTG 533  
|||  
DB 421 ATGACCCCTGGGCTGACTTCTTGAGACCACTGGAGTCTTTATCCTTGATTCATCTGTG 480  
OY 534 TTACTTTACCTTTACCTTTCTGTAGGCCCCAGAAAATTTATCAGCTTTTGTGAAAT 593  
|||  
DB 481 TTACTTTACCTTTACCTTTCTGTAGGCCCCAGAAAATTTATCAGCTTTTGTGAAAT 539  
OY 594 CTGGGCTGTCTCAAACTTGCCCTGTGCAGATACCCATCATGAGAAACATGGTCTTGGC 653  
|||  
DB 540 CTGGGCTGTCTCAAACTTGCCCTGTGCAGATACCCATCATGAGAAACATGGTCTTGGC 599

OY	654	CGAGCAATTCTGGGCTGGTGGACCCCTGTCCCAATTTGTAAGTTTCATATATGTGCAT	713
Db	600	CGGAGCAATTTCTGGGCTGGTGGACCCTTGTCACAATTTGTAAGTTTCATATATGTGCAT	659
OY	714	CCTCTGTGCTATCCTTCAGATCCAAITCAAAGGGAAGTTCAGAGGAAAAGCCTTCTGCACCTG	773
Db	660	CCTCTGTGCTATCCTTCAGATCCAAITCAAAGGGAAGTTCAGAGGAAAAGCCTTCCGCACCTG	719
OY	774	CTTCTCCACCTCTGTGTGATTTGAGACTCTTTATGCGCACAGCCATTTATCATGTATGTTGG	833
Db	720	CTTCTCCACCTCTGTGTGATTTGAGACTCTTTATGCGCACAGCCATTTATCATGTATGTTGG	779
OY	834	ACCAGATATGGGAACCCCAAGAGAGAGAANAATATCTCTGCTGTTTTACAGCCCTCT	893
Db	780	ACCAGATATGGGAACCCCAAGAGAGAGAANAATATCTCTGCTGTTTTACAGCCCTCT	839
OY	894	TAAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGAACCTCAGAAAGTGAAGAAATACTTT	953
Db	840	TAAATCCCATGCTCAATCCCCTTATCTGTAGTCTTAGGAACCTCAGAAAGTGAAGAAATACTTT	899
OY	954	GAAAGAGATGCTGGGAGTAGAAGGCGCTTTA	984
Db	900	GAAAGAGATGCTGGGAGTAGAAGGCGCTTTA	930

**RESULT 10**

ABA45207/c  
ID ABA45207 standard; DNA; 1957 BP.

AA ABA45207;  
AC

DT 01-FEB-2002 (first entry)

DE Human breast cell single exon nucleic acid probe #3902.

KM Human; microarray; single exon probe; gene expression; breast; disease; cancer; ss.

05 Homo sapiens.

PN W0200157271-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00662.

PR 04-FEB-2000; 2000US-0180312.

PR 30-JUN-2000; 2000US-0608408.

PR 21-SEP-2000; 2000US-0234687.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-496933/54.

PT	New spatially-addressable set of single exon nucleic acid probes,
PT	useful for measuring gene expression in sample derived from human
PT	breast, comprises number of single exon nucleic acid probes
XX	
PS	Claim 1; SEQ ID NO 3902; 327pp + sequence listing; English.

cc The invention relates to a spatially-addressable set of sim

CC probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for

CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a single exon nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from Wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 1957 BP; 604 A; 368 C; 495 G; 490 T; 0 other;

Query Match	53.78;	Score 545;	DB 22;	Length 1957;
Best Local Similarity	74.68;	Pred. No. 6e-153;		
Matches 711; Conservative	0;	Mismatches 240;	Indels 2;	Gaps 2

QY      36 TTTCCCTCTGTCATAGGATATGGGGACAAATATAACATCATCACAGAGTTCCTCCTACT    95  
||| || | ||| ||||| ||||||| ||| ||| ||||||| ||||||| |||||  
Db    1073 TTTTCTTTTTCACAGGGAATGGGGAAAAATCAGACAATGTTCACAGAGTTCCTCCTACT    1014

QY      96 GGGATTCCCGTTGCCCAAGATCAGATGCTCCTCTTGGGCTCTTCTCCCGTTCTTA 155  
|||||  
|||  
Db    1013 GGATTCCTCTGGGGCCAAGATCAGATGCTCCTCTTGGGCTCTTCTCCCTGTCTTA 954

QY 156 CGTCCTCACCCCTGCTGGGGACGGGACCATACTGGGGCTCATCTCACTGGACTTCAGACT 215  
|||||  
Db 953 TATCTTCACTGCTGTGGGACGGGACCATCTGGGGCTATCTCACTGGACTTCAGACT 894

QY 216 GCACGCCCC-TGTACTTCTCCTCACACACTGGCGGTGCATCGCCCTAGCCCTG 274  
||| ||||| |||||||||||||||| |  
Db 893 CCACACCCCCTACTCTTCCTTCACACCTGGTGCTGGCATCGCCCTACACCG 834

oy 275 CAACACGGTGGCCCGGATGCTGTGTAACCTCTGCATCAAGCCAAGCCCATCTCTTTGC 334  
 |||||  
 db 833 CAACACGGTGGCCCGGATGCTGTGTAACCTCTGCATCAAGCCAAGCCCATCTCTTTGC 774

335 GGGCGCATGTCAGACCTTCTGTTTTCACCTTTCGCTGTCAGAGATGTCCTCT 394  
 11 111111 1111111111 1 11 11111 11 1111111111 11  
 773 TCCCTTCAATACCTCCCAACCTCTTCTCTCTCTTCTCTCTTCTCTCTCTCTCT 714

[illegible]

455 CATGACCTGGAGAGTCTGCATCACCCTGGCGTGACTTCTTGAGACCACTGGAGTTCCTTTT 514

515 ATCCTGATTCATCTGTGTGTTACTTCTACCTTACCCCTCTGTAGGCCCCAGAAATTTA 574

575 TCACCTTTTGTGAAATCTTGGCTGTCTCAAACTTGGCTGTGCAGATACCCACATCA 634

DG 555 CCAC ACCCCTGGTATCCCTGTCCTCAAGGCGGCCTGTTGCTGACCTGGGTCTG 475

QY 635 ATGAGAACATGCTCTTGCCCGGAGCAATTCTGGGCTGTGGGACCCTTGTCCACAATTG 694

DD 4/4 ACCAGGAGGAGATCTTGCACCTTCGCGTCTCTCCGAGGGCCACCCACGCTGGGAGC 413  
 QY 695 TAGTTCATATATGTCATCCCTCTGCTATCCTTCAGATCCATCAAGGAGTTCAGA 754

DB 414 TTGTCTCTCACTGGACATCTCTGGGGCCATCTCTGAGCATCCAGTCTGGGGAGGGCCGCA 355

OY 755 GGAAAGCCCTTGCACCTGCTTCTTCCACCTCTGTGTGATTGGACTCTTTTATGGCACAG 814

Db 354 GAAAGGCCCTTCACACCTGCTCCCTCCACACCTCTGCGTGTGGAGCTCTTCTTTGCCAGTG 295

Oy 815 CCATTATCATGTATGTTGGACCCAGATATGGGAACCCCAAGAGCAGAGAATAATCTCC 874



Db 294 CCATCATCATGTACATGCCCCCAAGTCCGCCCATCTCTGAGAGACGAAAGGCTCTTTT 235  
QY 875 TGCCTGTTACAGCCCTCTTTAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAACT 934  
Db 234 TTCTATTTCACAGTTTTCACACCCCAACACTTAACCCCTGATTTACAGCCTGAGGAACG 175  
QY 935 CAGAGTGAAGATACTTTGAAGAGAGTGCCTGGAGTAGAAAGGGCTTTATGA 987  
Db 174 GAGAGTCAAGGGTCCCTGAGAGAGAGCACTGGGCAAGAAAGTCAATCTCTAA 122

RESULT 11  
ABA55696/c  
ID ABA55696 standard; DNA; 1957 BP.

AC ABA55696;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #4001.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.

PN MO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLB-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human foetal liver -

PS Claim 1; SEQ ID NO 4001; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a single exon nucleic acid

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pot\_sequences.

XX SQ Sequence 1957 BP; 604 A; 368 C; 495 G; 490 T; 0 other;

Query Match 53.7%; Score 545; DB 22; Length 1957;

Best Local Similarity 74.6%; Pred. No. 6e-153;

Matches 711; Conservative 0; Mismatches 240; Indels 2; Gaps 2;

QY 36 TTTCCTCTGTCATAGGATATGGGGACATATATCATCATCAGAGTCTCTCTACT 95

Db 1073 TTTCCTTTTTCACAGGGAATGGGGAAATCAGACAATGTCACAGACTCTCTACT 1014

QY 96 GGGATTTCCCGTTGGCCCAAGCATTCAGATGCTCTCTTTGGGCTCTCTCTCTCTA 155

Db 1013 GGGATTTCTCTCGGGCCCAAGCATTCAGATGCTCTCTTTGGGCTCTCTCTCTCTA 954

QY 156 CGTCTTACCCCTGCTGGGGAACGGACCATACTGGGGCTCATCTCATGACCTCAGACT 215  
Db 953 TATCTTACCCCTGCTGGGGAACGGGCAATCCTGGGGCTCATCTCATGACCTCAGACT 894  
QY 216 GCACGCCCCC-TGTAATCTCTCTCTCTCAACACTGGCGGCTCTGACATCGCCCTAGCCTG 274  
Db 893 CCACACCCCATGTAATCTCTCTCTCTCAACACTGGCTGTCTGACATCGCCCTAGACCG 834  
QY 275 CAACACGCTGCCCGGATGCTGTGAACCTCCTGCATCCAGCCCAAGCCCATCTCTTGGC 334  
Db 833 CAACACGCTGCCCGGATGCTGTGAACCTCCTGCATCCAGCCCAAGCCCATCTCTTGGC 774  
QY 335 GGGCCGATGATCAGACACTTCTGTTTCCACTTTTGGCTGTCAAGATGTCTCTCT 394  
Db 773 TGGTTCATGACGAGACCTTCTCTGTTTGAATTTGGACACAGCAATGTCTCTCT 714  
QY 395 GGTGATGATGCTATGATCTGTAGTGGCCATCTGCCACCCCTCCGATATTTGGCCAT 454  
Db 713 GGTGATGATGCTATGATCTGTAGTGGCCATCTGCCACCCCTCCGATATCTCTCT 654  
QY 455 CATGACCTGAGAGCTGTCATCACCCTGCGGTGACTTCTCTGGAACCACTGAGAGTCTTT 514  
Db 653 CATGACCTGAGAGCTGTCATCACCCTGCGGTGACTTCTCTGGAACCACTGAGAGTCTTT 594  
QY 515 ATCCTTGATTCATCTTCTGTACTTCTTACCTTTACCCCTCTGTAGGCCCCAGAAATTTA 574  
Db 593 GCTCTGAGCCCATGCTGTCTCTATCTCTAAGACTGCCCCCTCTGTGGGCTCATGAATCAA 534  
QY 575 TCACCTTTTTCGAATCTTGGCTGTCTCAACTTGGCTGTGACATACCCACATCA 634  
Db 533 CCAC-TTCTTCTGTGAATCTCTGTCTCTCTCAAGCTGGCTGTGACACCTGGCTCA 475  
QY 635 ATGAGAATGCTGTGGCCCGAGCAATTTCTGGGCTGTGGGACCCCTGTCCACAAATG 694  
Db 474 ACCAGTGTCTGTCTTGGAGCCCTGCTGTCTCTCTGTGGGACCCAGAGCTGTGC 415  
QY 695 TAGTTTCATATATGTCATCTCTGTCTCTCTCAACTTGGCTGTGACATACCAAGGAGTT 754  
Db 414 TTGCTCTCTACTGTCACATCTCTGGGCGCCATCTGAGGATTCAGTCTGGGAGGCGCA 355  
QY 755 GGAAGCCTTCTGACCTGCTCTGCCACCTCTGTGTGATTTGACCTTTTATGGACAG 814  
Db 354 GAAAGCCTTCTGACCTGCTCTGCCACCTCTGTGTGATTTGACCTTTTATGGACAGT 295  
QY 815 CCATATCATGTAATGTGGAGCCAGATATGGGAACCCCAAGAGAGCAGAAAGATATCTCC 874  
Db 294 CCATCATCATGTAATGTGGAGCCAGATATGGGAACCCCAAGAGAGCAGAAAGATATCT 235  
QY 875 TGCCTGTTACAGCCCTTTAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAACT 934  
Db 234 TTCTATTTCACAGTTTTCACACCCCAACACTTAACCCCTGATTTACAGCCTGAGGACG 175  
QY 935 CAGAGTGAAGATACTTTGAAGAGAGTCTGGAGTAGAAGGCTTTATGA 987  
Db 174 GAGAGTCAAGGGTCCCTGAGAGAGAGCACTGGGCAAGAAAGTCAATCTCTAA 122

RESULT 12

AAI13969/c  
ID AAI13969 standard; DNA; 1957 BP.

XX AAI13969;

DT 12-OCT-2001 (first entry)

DE Probe #3902 for gene expression analysis in human cervical cell sample.

KW Probe; human; microarray; gene expression; cervical epithelial cell;

OS Homo sapiens.

XX



PN WO200157278-A2.  
XX  
PD 09-AUG-2001.  
XX  
PE 30-JAN-2001; 2001WO-US00670.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
DR  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX  
PS Claim 25; SEQ ID No 3902; 487pp; English.  
XX  
CC The present invention relates to human single exon nucleic acid probes  
CC (SENP). The present sequence is one such probe. The SENPs are derived  
CC from human HeLa cells. The SENPs can be used to produce a single exon  
CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging  
CC of diseases of the cervix, notably cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1957 BP; 604 A; 368 C; 495 G; 490 T; 0 other;

Query Match 53.7%; Score 545; DB 22; Length 1957;  
Best Local Similarity 74.6%; Pred. No. 6e-153;  
Matches 711; Conservative 0; Mismatches 240; Indels 2; Gaps 2;

OY 36 TTTCCTCTGTATAGGATATGGGAGACATATACATCCATCAGAGTTCCTCTACT 95  
Db 1073 TTTCCTTTTTCACAGGAAATGGGAAATCAGACATGTCACAGAGTTCCTCTACT 1014  
OY 96 GGGATTTCCCGTGGCCCAAGGATTCAGATGCTCTCTTGGGCTCTTCTCCCTGTCTA 155  
Db 1013 GGGATTTCTCTGGGCCCAAGGATTCAGATGCTCTCTTGGGCTCTTCTCCCTGTCTA 954  
OY 156 CGTCTTCACCCCTGCTGGGGAACGGGACCATCTAGGGCTCATCTCAGTGCAGACT 215  
Db 953 TATCTTACCCCTGCTGGGGAACGGGACCATCTAGGGCTCATCTCAGTGCAGACT 894  
OY 216 GCAGGCCCCC-TGTACTTCTCTCTCAACCTGGCGTGTGACATCGCCTACGCCTG 274  
Db 893 CCACACCCCATGTACTTCTCTCTCAACCTGGCGTGTGACATCGCCTACGCCTG 834  
OY 275 CAACAGGTCGCCCGGATGCTGTGAACCTCTGATCAGTCAAGCCCAAGCCATCTCTTGC 334  
Db 833 CAACAGGTCGCCCGGATGCTGTGAACCTCTGATCAGTCAAGCCCAAGCCATCTCTTGC 774  
OY 335 GGGCCGATGATGCAGACCTTCTGTTTCCACTTGTGCTGTCAAGAAATGTCCTCTCT 394  
Db 773 TGGTTGATGACGACACCTTCTCTGTTTGAAGTTTGGACACAGCGAATGTCCTCTCT 714  
OY 395 GGTGGTATGTCCTATATCTGTAGCTGGCCATCTGCCACCCCTCCGATATTTGGCCAT 454  
Db 713 GGTGGTATGTCCTATATCTGTAGCTGGCCATCTGCCACCCCTCCGATATTTGGCCAT 654  
OY 455 CATGACCTGAGAGTGTGATCACCCTGCGGGTGACTTCTGTGAGACCACTGAGTCCCTTT 514  
Db 653 CATGACCTGAGAGTGTGATCACCCTGCGGGTGACTTCTGTGAGACCACTGAGTCCCTCT 594

OY 515 ATCCCTGATTCATCTTGTGTACTTCTACCTTTACCTTCTGTAGGCCACGAAATTTA 574  
Db 593 GGCCTGCCCCATGTGTGTCTATCTTACAGCTGCTTCTCTGGGCTCATGAAATCAA 534  
OY 575 TCACCTTTTGTGAAATCTGGCTGTCTCAAACTTGCCCTGTGAGATFACCCACATCA 634  
Db 533 CCAC-TTCTTCTGTGAATCTGTCTGTCTCCTCAGCTGGCCTGTGCTGACACCTGGCTCA 475  
OY 635 ATGAGACATGCTCTGGCCGAGCAATTTCTGGGCTGGTGGACCTTGTCCACAATTG 694  
Db 474 ACCAGTGGTCATCTTTGACAGCTGCGTGTCTTCTCTGTTGGGCCACCCAGCTGTGTC 415  
OY 695 TAGTTTCAATATATGTGATCTCTGTGCTATCTTCAGATCCAAATCAAGGAATTCAGA 754  
Db 414 TTGTCTCTACTCGCACATCTGGCGGCGCATCTGAGGATCCAGTCTGGGAGGCGCCGA 355  
OY 755 GGAAGCCTTCTGCACCTGCTTCCACCTCTGTGTGATTTGACTCTTTATGGCACAG 814  
Db 354 GAAAGCCTTCTCCACCTGCTTCCACCTCTGTGTGATTTGACTCTTTATGGCACAG 295  
OY 815 CCATTATCATGTATGTGACCCAGATATGGAAACCCCAAGAGACAGAAATATCTCC 874  
Db 294 CCATCATCATGTATGATGGCCCCCAAGTCCCGCCATCTGAGGAGACGAAAGTCTTTT 235  
OY 875 TGCTGTTCACAGCCTCTTTAAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAACT 934  
Db 234 TTCTATTATTACAGTTTTCACACCCCAACACTTAACCCCTGATTTACAGCCTGAGGACG 175  
OY 935 CAGAGTGAAGATACTTTGAAGAGAGTGTGGGAGTAGAAAGGCTTTATGA 987  
Db 174 GAGAGGTCAAGGCTGCCCCGAGAGAGACACTGGGCAAGGAAGTCAATTCCTTA 122

RESULT 13  
AAI35352/C  
ID AAI35352 standard; DNA; 1957 BP.  
XX  
AC AAI35352;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Probe #4038 used to measure gene expression in human placenta sample.  
XX  
KW Probe; microarray; human; placenta; antenatal diagnosis;  
KW genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00663.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
DR  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
PS Claim 25; SEQ ID No 4038; 654pp; English.





XX The sequence encodes a human G-protein coupled receptor (GPCR)-like  
CC protein. The GPCR-like polypeptides and polynucleotides are useful for  
CC the treatment of diseases of ophthalmic, neurological (e.g Alzheimer's  
CC disease and Parkinson's disease, immunological (e.g HIV infection and  
CC candidiasis), autoimmune disorders (e.g multiple sclerosis, systemic  
CC lupus erythematosus and rheumatoid arthritis), platelet disorders (e.g  
CC thrombocytopenia and aplastic anaemia), inflammatory disorders (e.g.  
CC septic shock and systemic inflammatory response syndrome, SIRS) and  
CC nephritic systems. They may also be used to treat hormonal dysfunction,  
CC cancer, atherosclerosis, wounds, psoriasis and diabetes. Numerous examples  
CC of each type of disorder are given in the specification. Anti-GPCR-like  
CC protein antibodies are useful for detecting or quantitating the  
CC polypeptide in tissue. The polypeptides can also be used as molecular  
CC weight markers and as a food supplement.

SQ Sequence 2282 BP; 583 A; 573 C; 428 G; 698 T; 0 other;

Query Match 53.1%; Score 538.2; DB 22; Length 2282;  
Best Local Similarity 73.6%; Pred. No. 7.2e-151;  
Matches 712; Conservative 0; Mismatches 253; Indels 2; Gaps 2;

QY 31 CTGATTTCTCTGTGATAGGATATGGGGACAAATATACATCCATCAGAGTTCCTC 90  
DB 1233 CTTTGTGTTTGTAGTGAATGTGAATAATCAGACAATGTGCACAGAGTTCCTC 1292

QY 91 CTAAGGATTTCCCGTGGCCCAAGGATGATGCTCCTTTGGGCTCTTCTCCCTG 150  
DB 1293 CTAAGGATTTCTCTGGGCCCCAAGGATGATGCTCCTTTGGGCTCTTCTCCCTG 1352

QY 151 TTCTACGTTCTTCAACCTGCTGGGGAACGGACCATACTGGGCTCATCTCACTGACTCC 210  
DB 1353 TTCTATGTTCTTCAACCTGCTGGGGAATGGACCATCTGGGCTCATCTCACTGACTCC 1412

QY 211 AGACTGACGCCCCCTGACTTCTCTCTGACACCTGGGCTGCTGACATCGCCTAC 269  
DB 1413 AGACTGACACCCCCCATGACTTCTCTCTGACACCTGGGCTGCTGACATCGCCTAT 1472

QY 270 GCGTGAACACAGGTGCGCCGATGCTGTGAACTCTGTCATCCAGCCAGCCCATCTCC 329  
DB 1473 GCGTGAACACAGGTGCGCCGATGCTGTGAACTCTGTCATCCAGCCAGCCCATCTCC 1532

QY 330 TTTGGGGGCGCATGATGACACCTTCTGTTTCCACTTTTGCTGTACAGATGCTC 389  
DB 1533 TTTGCTGGCTGATGACACACCTTCTGTTTGAAGTTTGCACATGATGCTC 1592

QY 390 CTCCGTGGTGTGATGCTGATGATCTGTAGTGGCCATCTGCCACCCCTCCGATATTG 449  
DB 1593 CTCCGTGGTGTGATGCTGATGATCTGTAGTGGCCATCTGCCACCCCTCCGATATTG 1652

QY 450 GGCATCATGACCTGAGAGTCTGATCACCCTGCGGTGACTTCCCTGACCACTGAGTC 509  
DB 1653 ATCATCATGACCTGAGAGTCTGATCACCCTGCGGTGACTTCCCTGACATGTGCTCC 1712

QY 510 CTTTATTCCTTGATTCATCTGTGTACTTCTTACCTTTACCCTTCTGTAGGCCCCAGAA 569  
DB 1713 CTCCGTGGCTATGCTCATGAGACCTCATGCTAAGACTGCCCCCTTTGTGGGCTCTGAA 1772

QY 570 ATTATGACTTTTGTGAAATCTTGCTGCTTCAAACTTGCTGCTGACAGATACCCA 629  
DB 1773 ATCAACAC-TTCTTCTGTAATCTGCTGCTCAGGCTGGCCTGCTGATACCTG 1831

QY 630 CATCATGAGACATGCTTGGCCGAGCAATTTCTGGGCTGCTGGACCTTGTCCAC 689  
DB 1832 GCTCAACCAAGTGTGCTTGTGACGCTCATGCTTCACTCTGCTGGACCACTTGCCT 1891

QY 690 AATTGTAGTTTCATATATGTCATCTCTGTGCTATCTTCAATCAAGGAAGT 749  
DB 1892 GGTGCTGCTCTCTACTCATCATCTCTGGGCACTCTGAGGATCCAGTCTGGGAGGG 1951

QY 750 TCAGAGGAAGCCTCTGCACTGCTCTCCCACTCTGTGATTTGACTCTTTATGG 809  
DB 750 TCAGAGGAAGCCTCTGCACTGCTCTCTCCCACTCTGTGATTTGACTCTTTATGG 809

DB 1952 CCGAGAAAGCCCTTCTCCACCTGCTCTCTCCACCTCTGCGTAGTGGACTCTTTTGG 2011  
QY 810 CACAGCCATTTATGATGATGTTGGACCAAGATATGGAAACCCCAAGGACAGAAATA 869  
DB 2012 CAGGCCATGCTATGATGATGAGCCCTTAAGTCCCGCCATCTGAGGAGCAGAGAGT 2071

QY 870 TCTCCTGCTGTTTCAACAGCCCTTTAATCCCATGCTCAATCCCTTATCTGATCTTAG 929  
DB 2072 CTTTCTTATTTTACAGTCTTTCAACCCGATGCTTAACCCCTGATTTACACCTGAG 2131

QY 930 GAACCTCAGAAGTGAAGATCTTTGAAGAGAGTGTGAGAGTGAAGAGGCTTTATGAAA 989  
DB 2132 GAATGTAGAGTCAAGGGTGCCTGAGAGAGCACTGTGCAAGGAAAGTCAATTCCTAAGA 2191

QY 990 AGGATTA 996  
DB 2192 GGTGTGA 2198

Search completed: May 2, 2003, 05:01:21  
Job time : 293 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 04:52:26 ; Search time 1543 seconds  
(without alignments)  
10643.048 Million cell updates/sec

Title: US-09-898-586-23  
Perfect score: 1014  
Sequence: 1 taacactctctcctaacca.....tatggcattgtgactgacaa 1014

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1:	em_estba:*
2:	em_esthum:*
3:	em_estlin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_hic:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_hic:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_luv:*
20:	em_gss_pln:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_mam:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_frod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	532.8	52.5	2021	11 BC016940	BC016940 Homo sapi
2	432	42.6	456	9 AI339565	AI339565 qk67e07.x
3	429.4	42.3	443	13 BG984967	BG984967 IL5-CN006
4	417.4	41.2	456	14 BQ301355	BQ301355 QV1-HB003
5	409	40.3	433	9 AI339573	AI339573 qk67f03.x
6	398.4	39.3	619	9 AI148854	AI148854 qct70a02.x

7	394.8	38.9	705	11 AF327904	AF327904 Homo sapi
8	394.8	38.9	751	12 BE856570	BE856570 T64e08.x
9	380.6	37.5	738	9 AI079550	AI079550 0204f08.x
10	370.4	36.5	732	12 BF116115	BF116115 7n77g04.x
11	359	35.4	580	10 BE543056	BE543056 601068985
12	355	35.0	367	9 AI340119	AI340119 qk64f08.x
13	345.6	34.1	426	14 H62441	H62441 Yr79d08.r1
14	345.6	34.1	623	13 BM670360	BM670360 UI-E-DW1-
15	337.4	33.3	361	9 AI285267	AI285267 q138c08.x
16	332.6	32.8	690	10 BE179276	BE179276 RCI-HT061
17	331	32.6	439	14 H62445	H62445 Yr79e08.r1
18	320	31.6	441	13 BI020419	BI020419 CM3-MT029
19	305.2	30.1	499	10 AW299289	AW299289 XS38f05.x
20	293.8	29.0	332	10 BE182734	BE182734 RC3-HT065
21	272.4	26.9	788	17 BH111304	BH111304 RPCI-24-3
22	267.4	26.4	432	10 BE174029	BE174029 QV1-HT057
23	265.2	26.2	671	12 BF974217	BF974217 602243851
24	242.8	23.9	520	17 AO600752	AO600752 HS_5346.B
25	232	22.9	376	12 BE926655	BE926655 QV1-BT063
26	222.6	22.0	449	14 H39851	H39851 YP01h11.r1
27	214.4	21.1	432	9 AA962054	AA962054 Oq79901.s
28	211.8	20.9	760	12 BG169291	BG169291 602321020
29	206.8	20.4	1394	11 AK017036	AK017036 Mus muscu
30	206.8	20.4	3063	11 AK016560	AK016560 Mus muscu
31	203	20.0	445	14 H39853	H39853 YP01h12.s1
32	183.6	18.1	226	9 AI084964	AI084964 Ow86f03.s
33	180	17.8	231	13 BG984943	BG984943 IL5-CN006
34	174.4	17.2	796	12 BG197640	BG197640 RST17016
35	174	17.2	762	12 BG193339	BG193339 RST12467
36	173.6	17.1	495	12 BF841098	BF841098 PM2-HT035
37	173.4	17.1	661	13 BG924888	BG924888 HNC36-1-B
38	162.8	16.1	445	14 BM712072	BM712072 UI-E-DW1-
39	162.4	16.0	428	14 BM707299	BM707299 UI-E-CRI-
40	160.8	15.9	642	17 AZ969227	AZ969227 2M0241J24
41	160.4	15.8	1501	11 AK016338	AK016338 Mus muscu
42	160.2	15.8	732	12 BG068751	BG068751 H3068H10-
43	156.4	15.4	786	12 BG204873	BG204873 RST24289
44	155.6	15.3	915	14 BQ888002	BQ888002 AGENCOURT
45	153.8	15.2	632	17 AZ765752	AZ765752 1M0562E19

## ALIGNMENTS

RESULT 1  
LOCUS BC016940 2021 bp mRNA linear HTC 09-NOV-2001  
DEFINITION Homo sapiens, similar to olfactory receptor, family 2, subfamily A, member 4, clone IMAGE:4424116, mRNA.

ACCESSION BC016940  
VERSION BC016940.1 GI:16877381

KEYWORDS HTC.  
SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 2021)  
AUTHORS Strausberg, R.

TITLE Direct Submission  
JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcdepaxll.stanford.edu](mailto:mcdepaxll.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: TRAK Plate: 27 Row: 1 Column: 22

Series: 27 Row: 1 Column: 22  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction, Similarity but not identity to protein

This clone has the following problem: frame shifted to protein

FEATURES	Location/Qualifiers
source	1. .2021

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4424116"
/tissue_type="Kidney, hypernephroma"
/clone_1lb="NH_MGC_89"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

```

BASE COUNT	470 a	546 c	427 g	578 t
ORIGIN				

Query Match	52.5%	Score 532.8;	DB 11;	Length 2021;
Best Local Similarity	73.5%;	Pred. No. 1.9e-134;		
Matches 706; Conservative	0;	Mismatches 252;	Indels 2;	Gaps 2;

OY		38	TCCCTCTGTCATAGGGAATAAGGGGCACAAATATACATCATTCAACAGAGTTTCTCTACTGG	97
Db		412	TCATTAATAATTAGTGAAAATGGTGAAAAAATCAGACAATGGTCACAGAGTTTCTCTACTGG	471
OY		98	GATTTCGCCGTGGCCCCAAGGATTACAGATGCTCTCTTTGGGGCTCTCTCCCTGTCTACG	157
Db		472	GATTTCCTCTGGGGCCCNAAGGATTACAGATGCTCTCTTTGGGGCTCTCTCCCTGTCTATG	531
OY		158	TCCTCACCCCTGCTGGGGAACCGGACCATACTGGGGCTCATCTCAGCTGGACTCCAGACTGC	217
Db		532	TCCTCACCCCTGCTGGGGAAATGGGACCATCCCTGGGGCTCATCTCAGCTGGACTCCAGACTGC	591
OY		218	ACGCCCC-TGTA CTCTCTCTCTC ACAC C T G CGG GT C G TG A CA AT CG CCT AG CC TT GA	276
Db		592	ACACCCCCATGTA CTCTCTCTCTC ACAC C T G CGG GT C G TC AA CA AT CG CCT AG CC TT GA	651
OY		277	ACACGGTGGCCCCGATGCTGSTGAACCTCTGCATCCAGCCAAGCCCATCTCTTGGCGG	336
Db		652	ACACAGTGGCCCCAGATGCTGSTGAACCTCTGCATCCAGCCAAGCCCATCTCTTGGCTG	711
OY		337	GCCGCATGATGCAGACCTTTCTGTTTTCCACTTTTGGCTGTACAGAATGTCCTCTCTGG	396
Db		712	GCTGCATGATGCATAGACCTTTCTCTTTTGAGTTTGGACATACTGAATGCTCTCTGTGG	771
OY		397	TGGTGATGTCCTAATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCATCA	456
Db		772	TGCTGATGTCCTACGATCGGTACGTGGCCCATCTGCCACCCCTCTCCGATATTGCATCA	831
OY		457	TGACCTGGAGAGTCTGCATCAACCTCGCGGTGACTTCCTGGACC ACT GG AG TTC TTT AT	516
Db		832	TGACCTGGAAAGTCTGCATCACTCTGGCCCATCACTTCCTGGACATGTGGCTCCCTCTGG	891
OY		517	CCTTGATTCATCTTGTGTTACTTCTACCTTACCCTTCGTAGGCCCCAGAAAATTATC	576
Db		892	CTATGCTCATGTGAGCCCTCATCTCTAAGACTGCCCTTTTGGGGCTCGTGAAATCAAC	951
OY		577	ACTTTTTTGTGAATCTTGGCTGTCTCAAACCTTGCCCTGTSCAGATA ACC CATCAAT	636
Db		952	AC- TTCTTCTGTGAATCTGTCTGTCTCAGGCTGGCCGTGTGATACCTGGCTCAAC	1011
OY		637	GAGAACAATGTCCTGGCCCGAGCAATTTCTGGGCTGTGGGAGCCCTGTCCACAATGTA	696
Db		1011	CAGGTGTCATCTTTGCAGCCCTGCATGTTCACTCCTGGTGGGACCACACTCTGCCCTGGTCTG	1077
OY		697	GTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGAGTTCAGAGS	756

Db	1071	GTCCTACTCACACATCCTGGCGGCATCCTGAGCATCCAGTCTGGGAGGGCCGCAGA	1130
QY	757	AAAGCCTTCTGCACCTGCTTCGCCACCTCTGTGTGATTGGACTCTTTATGGCACAGCC	816
Db	1131	AAGCCTTCTCCACCTGCTCTCCACCTCTGCGTAGTGGAACTCTTCTTTGGCAGGCC	1190
QY	817	ATTATCATGTATGTTGACCCAGATATGGGAACCCCAAGAGCAGAGAATAATCTCCTG	876
Db	1191	ATGCTCATGTACATGGCCCCCTAACTCCCGCCATCTCTGAGCAGCAGCAGAAGTCCCTTTT	1250
QY	877	CTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCCCTTATCTGTAGTCTTAGGAATCA	936
Db	1251	CTATTTTACAGTCTCTTTCACCCGATGCTAAACCCCTGATTTACAACCTGAGGAATGTA	1310
QY	937	GAACTGAAGAATACTTTGAAGAGAGCTGCTGGAGTAGAAGGGCTTATGAAAAGGATTA	996
Db	1311	GAGGTCAAGGGTGGCCCTGAGCAGAGCACTGTGCAAGAAAGTCAATTCTTAAGAGGTGTA	1370

RESULT 2	AI339565/c	LOCUS	DEFINITION
AI339565	456 bp	mRNA	linear
qk67e07.x1	NCI CGAP C08 Homo sapiens CDNA clone IMAGE:1874052 3'		
similar to SW:OLF7_MOUSE P34984	OLFACTORY RECEPTOR-LIKE PROTEIN K7		
;', mRNA sequence.			

FEATURES	Location/Qualifiers
source	1. .456

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1874052"
/clone_1lb="NCI_CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

```

BASE COUNT	136 a	93 c	149 g	78 t
ORIGIN				

Query Match	42.68;	Score 432;	DB 9;	Length 456;
Best Local Similarity	99.88;	Pred. No. 3.7e-107;		
Matches 443; Conservative	0;	Mismatches 0;	Indels 1;	Gaps 1;

143 TCTCCCTGTTCTACGTCCTTCAACCTGCTGGGGAACGGACCACTACTGGGGCTCATCTCAC 202

```

|||||
Db 456 TCTCCTGTTCAGTCTTCACCCCTGGGGGAACGGACCACTACTGGGCTCATCTCAC 397
QY 203 TGGACTCCAGACTGACACGCCCCC-TGACTCTCTCTCTCACACCTGGCGTCTGACA 261
Db 396 TGGACTCCAGACTGACACGCCCCCAGTACTCTCTCTCACACCTGGCGTCTGACA 337
QY 262 TGGCTACGCTGACACACGGTGGCCGAGTGTGTGAACCTCTGATCCAGCAAGC 321
Db 336 TGGCTACGCTGACACACGGTGGCCGAGTGTGTGAACCTCTGATCCAGCAAGC 277
QY 322 CCATCTCCTTTGGGGCCGATGATGACACCTTCTGTTCACCTTTGCTGTACAG 381
Db 276 CCATCTCCTTTGGGGCCGATGATGACACCTTCTGTTCACCTTTGCTGTACAG 217
QY 382 AATGCTCCTCTGCTGTGTGATGCTCTGATCTGTACGTGGCCATCTGCCACCCCTCC 441
Db 216 AATGCTCCTCTGCTGTGTGATGCTCTGATCTGTACGTGGCCATCTGCCACCCCTCC 157
QY 442 GATATTTGGCCATGACCTGAGAGCTGCATCACCCCTGGGTGACTTCTGTGACCA 501
Db 156 GATATTTGGCCATGACCTGAGAGCTGCATCACCCCTGGGTGACTTCTGTGACCA 97
QY 502 CTGAGTCTCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTT 561
Db 96 CTGAGTCTCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTT 37
QY 562 CCCAGAAAATTATCACTTTT 585
Db 36 CCCAGAAAATTATCACTTTT 13

RESULT 3
BG984967/c 443 bp mRNA linear EST 12-JUN-2001
LOCUS IL5-CN0068-100401-434-d09 CN0068 Homo sapiens cDNA, mRNA sequence.
DEFINITION BG984967
ACCESSION BG984967
VERSION BG984967.1 GI:14387702
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 443)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&tl2=IL5-CN0068-
100401-434-d09&ts=2001-04-10&tl=1)
Seq primer: puc 18 forward
High quality sequence stop: 363.
FEATURES
source
1..443
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CN0068"

```

```

/dev_stage="Adult"
/Note="Organ: Colon normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 136 a 94 c 131 g 81 t 1 others
ORIGIN

```

```

Query Match 42.3%; Score 429.4; DB 13; Length 443;
Best Local Similarity 99.3%; Pred. No. 1.9e-106;
Matches 441; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

```

QY 261 ATGCGCTACGCTTGCAACACAGGTGCCCCGATGCTGTGAACCTCTGCATCCAGCCAAG 320
Db 443 ATGCGCTACGCTTGCAACACAGGTGCCCCGATGCTGTGAACCTCTGCATCCAGCCAAG 384
QY 321 CCCATCTCTTTGGCGGCGCATGATGACACCTTCTGTGTTTCCACTTTGCTGTGACA 380
Db 383 CCCATCTCTCTGCTGTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGT 324
QY 381 GAATGCTCTCTCTGCTGTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGT 440
Db 323 GAATGCTCTCTCTGCTGTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGTGATGCTGT 264
QY 441 CGATATTTGGCCATCATGACCTGAGAGTCTGCATCACCCCTGGGTGACTTCTGTGACC 500
Db 263 CGATATTTGGCCATCATGACCTGAGAGTCTGCATCACCCCTGGGTGACTTCTGTGACC 204
QY 501 ACTGAGTCTCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTT 560
Db 203 ACTGAGTCTCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTTATCTT 144
QY 561 CCCAGAAAATTATCACTTTTGTGAATCTTGGCTGTCTCAAACTTGCCTGTGC 620
Db 143 CCCAGAAAATTATCAC-TTTTGTGAATCTTGGCTGTCTCAAACTTGCCTGTGC 85
QY 621 AGATACCAATCATGAGAACATGCTTGGCCGAGCAATTCTGGCGTGTGGAGCC 680
Db 84 AGATACCAATCATGAGAACATGCTTGGCCGAGCAATTCTGGCGTGTGGAGCC 25
QY 681 CTGTCCACAATTGTAGTTTCATA 704
Db 24 CTGTCCACAATTGTAGTTTCATA 1

RESULT 4
BQ301355
LOCUS BQ301355 456 bp mRNA linear EST 16-MAY-2002
DEFINITION QV1-HB0037-020201-589-d09 HB0037 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ301355
VERSION BQ301355.1 GI:20816877
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 456)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

```





Db	13	T	13
RESULT 6			
LOCUS	AI148854		
DEFINITION	AI148854	619 bp	mRNA linear EST 28-OCT-1998
ACCESSION	gc70a02.x1 Soares.placenta_8to9weeks_2NHP8to9w		Homo sapiens CDNA
VERSION	clone IMAGE:1714922 3'		similar to SW:OLF7_MOUSE P34984 OLFATORY
KEYWORDS	RECEPTOR-LIKE PROTEIN K7 ;		mRNA sequence.
SOURCE	AI148854		
ORGANISM	AI148854.1	GI:3677323	
REFERENCE	EST.		
AUTHORS	human.		
TITLE	Homo sapiens		
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
COMMENT	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	1 (bases 1 to 619)		
	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .		
	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Tumor Gene Index		
	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: <a href="mailto:cgaps-remail.nih.gov">cgaps-remail.nih.gov</a>		
	This clone is available royalty-free through LNL ; contact the		
	IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further information.		
	Insert Length: 1098 Std Error: 0.00		
	Seq primer: -40m13 fwd. ET from Amersham		
	High quality sequence stop: 432.		
FEATURES	Location/Qualifiers		
SOURCE	1. 619		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:1714922"		
	/dev_stage="Soares.placenta_8to9weeks_2NHP8to9w"		
	/dev_stage="two placenta: one from 8 weeks and another		
	from 9 weeks post conception"		
	/lab_host="DH10B (ampicillin resistant)"		
	/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a		
	modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st		
	strand cDNA was primed with a Not I - oligo(dT) primer [5'		
	TGTTACCAATCTGAAGTGGAGCGCGCGGATTTTCTTTTCTTTT 3'] ,		
	double-stranded cDNA was size selected, ligated to Eco RI		
	adapters (Pharmacia), digested with Not I and cloned into		
	the Not I and Eco RI sites of a modified pT7T3 vector		
	(Pharmacia). Library constructed by Bento Soares and		
	M.Fatima Bonaldo."		
BASE COUNT	112 a	191 c	131 g 183 t 2 others
ORIGIN			
Query Match	39.3%;	Score 398.4;	DB 9; Length 619;
Best Local Similarity	80.3%;	Pred. No. 6.5e-98;	
Matches 490; Conservative	0;	Mismatches 118;	Indels 2; Gaps 2;
31 CTGTGATTCCTCTGTGTCATAGGAGATATGGGGACATAATATACATCCATCAGAGAGTTCTTC	90		
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111			
2 CTTTGTTCCTTTTGTAGTGAATAGTGTAATAATCAGACAATGTGCACAGAGTTCTTC	61		
91 CTAATGGGATTTCCCGTTGGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTTCTCCCTG	150		
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111			
62 CTAATGGGATTTCTCTCTGGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTTCTCCCTG	121		
151 TTCTAGCTCTTACCCCTGCTGGGGGAACGGGACCATACTTGGGGCTCATCTACTGAGCTCC	210		
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111			
122 TTCTATGTCTTACCCCTGCTGGGGGAATGGGACCATCTGCGGCTCATCTACTGAGCTCC	181		
211 AGACTGCAGCGCCCC -TGTAATTTCTTCTCTCAGACCTGGGGCTGTCGACATGGCTTAC	269		
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111			
182 AGACTGCAGACCCCAATGTAATCTTCTCTCAGACCTGGGCTGTCAGAGCTGGCTTAT	241		
270 GCCTGCAGACAGGTGCCCCGGGATGCTGTGAACCTCTCTGATCCAGCCAGCCCATCTCC	329		
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111			
242 GCCTGCAGACAGGTGCCCCGGGATGCTGTGAACCTCTCTGATCCAGCCAGCCCATCTCC	301		

[illegible]



```
Db 109 CTTTGTGTTTGTAGTAGTGAATGGTGAAAAATCAGACATGATGTCACAGAGTTCCCTC 168
QY 91 CTACTGGATTCCCGCTGGCCCAAGATTCAGATGCTCCTTTGGGCTCTCTCCCTG 150
Db 169 CTACTGGATTCTCTCTGGGCCCAAGATTCAGATGCTCTCTTTGGGCTCTCTCCCTG 228
QY 151 TTCTACGCTTTCACCTGCTGGGAGACGGGACCATATCGGGGCTCATCTCAGTGGACTCC 210
Db 229 TTCTATGCTTTCACCTGCTGGGAGATGGGACCATCTGGGGGCTCATCTCAGTGGACTCC 288
QY 211 AGACTGCACGCCCC-TGTACTTCTCTCTCTCAGACCTGGCGGCTGTCGACATCGCCCTAC 269
Db 289 AGACTGCACACCCCATGACTTCTCTCTCTCAGACCTGGCGGCTGTCGACATCGCCCTAT 348
QY 270 GCCTGCAACACGGTGGCCCGGATGCTGGTGAACCTCTGTCATCCAGCCCAAGCCCATCTCC 329
Db 349 GCCTGCAACACAGTGGCCCGGATGCTGGTGAACCTCTGTCATCCAGCCCAAGCCCATCTCC 408
QY 330 TTGCGGGGCGCATGATGACAGACCTTCTGTTTCACATTTTGTCTGTCACAGATGCTCTC 389
Db 409 TTTGCTGGCTGATGACATGACCTTCTCTTTTGTAGTTTGGACATATGATGCTCTC 468
QY 390 CTCTGCTGCTGATGCTCTATGATCTGATGCTGGCCATCTGCCACCCCTCCGATATTTC 449
Db 469 CTGTGCTGCTGATGCTCTACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
QY 450 GCCATCATGACCTGAGAGTCTGACATCACCTCGCGGTGACTTCTCTGACCACTGGAGTC 509
Db 529 ATCATCATGACCTGAGAGTCTGACATCACCTCTGACATCACCTGACATGCTGCTGCTGCT 588
QY 510 CTTTATCCTGATTCATCTTGTGTACTTCTTACCTTTCCTTCTGATGAGCCCAAGAA 569
Db 589 CTCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAA 648
QY 570 ATTTATCCTTTTGTGTAATCTTGGGCTGTTCTCAAACTTGCCCTGTGAGATACC 627
Db 649 ATCAACACAC-TTCTTCTGTGAAACCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705

RESULT 8
BE856570 751 bp mRNA linear EST 29-SEP-2000
LOCUS 764e08.x1 Soares_NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
DEFINITION IMAGE:3299462 3' similar to TR:095047 095047 WUGSC:H_DJ0988615.2
PROTEIN .; mRNA sequence.
ACCESSION BE856570
VERSION BE856570.1 GI:10369724
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 751)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 449.
FEATURES
Location/Qualifiers
1..751
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3299462"
/clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pTZ19-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
```

```
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and cloneIDs: Soares NbHSF pool 1:
309384-310919, 323208-325895 Soares Nb2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares Nb2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NbHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NbHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 170 a 206 c 137 g 238 t
ORIGIN
Query Match 38.9%; Score 394.8; DB 12; Length 751;
Best Local Similarity 80.9%; Pred. No. 6.9e-97;
Matches 484; Conservative 0; Mismatches 112; Indels 2; Gaps 2;

QY 31 CTGATTTCTCTCTGTCATAGGATATGGGGGACATATACATCCATCAGAGTCTCTC 90
Db 155 CTTTGTGTTTGTAGTAGTGAATGCTGAATAATCAGACATGCTCAGAGTCTCTC 214
QY 91 CTACTGGATTTCCTCGTGGCCCAAGATTGATGCTCTCTTTGGGCTCTCTCCCTG 150
Db 215 CTACTGGATTTCCTCGTGGCCCAAGATTGATGCTCTCTTTGGGCTCTCTCCCTG 274
QY 151 TTCTAGTCTTACCCCTGCTGGGGAACGGGACCATCTGAGGCTCTCTCTCTCTCTC 210
Db 275 TTCTATGCTTTCACCCCTGCTGGGGAATGGACCATCTCTGGGCTCTCTCTCTCTC 334
QY 211 AGACTGCACGCCCC-TGTACTTCTCTCTCTCAGACCTGGCGGTGCTGACATGCTTAC 269
Db 335 AGACTGCACACCCCATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 394
QY 270 GCCTGCAACACGGTGGCCCGGATGCTGTAACCTCTCTGATCCAGCCCAAGCCCATCTCC 329
Db 395 GCCTGCAACACAGTGGCCCGGATGCTGTAACCTCTCTGATCCAGCCCAAGCCCATCTCC 454
QY 330 TTTGCGGGCGCATGATGACAGACCTTCTGTCTTCTCACTTTTGTGCTGACAGATGCTCTC 389
Db 455 TTTGCTGCTGATGACATGACATGACCTTCTCTTTGAGTTTGGACATGATGATGCTCTC 514
QY 390 CTCTGCTGCTGATGCTCTATGATCTGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 449
Db 515 CTGTGCTGCTGATGCTCTATGATCTGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 574
QY 450 GCCATCATGACCTGAGAGTCTGATCACCCTCGCGGTGACTTCTCTGAGACCACTGGAGTC 509
Db 575 ATCATCATGACCTGAGAGTCTGATCACCCTCTTGGCATCACTTCTGAGACATGTGGCTTC 634
QY 510 CTTTATCCTGATTCATCTTGTGTACTTCTTACCTTTTACCCTTCTGTAGGCCCAAGAA 569
Db 635 CTCTGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 694
QY 570 ATTTATCCTTTTGTGTAATCTTGGCTGTCTCTCAAACTTGCCCTGTGAGATACC 627
Db 695 ATCAACACAC-TTCTTCTGTGAAACCTGTCTGCTCTCAGCTGGCCCTGTGATACC 751

RESULT 9
A1079550 738 bp mRNA linear EST 29-SEP-1998
LOCUS A1079550
DEFINITION oz04f08.x1 Soares_fetal_liver_spleen_1NF15_S1 Homo sapiens cDNA
clone IMAGE:1674375 3' similar to SW:OLP7_MOUSE P34984 OLFATORY
RECEPTOR-LIKE PROTEIN K7 ;, mRNA sequence.
ACCESSION A1079550
VERSION A1079550.1 GI:3415801
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



```

OY 215 TGCAGCCCCC-TGTACTTCTCTCTCTCACACCTGGCGGTGCTGACATCGCCTACGCCCT 273
Db 380 TCCACACCCCCCATCTCTCTCTCTCACACCTGGCTGTGCTGACATCGCCTACACCC 439
OY 274 GCACACGGTGGCCCCGATGCTGTGAACCTCTGATCCAGCCAGCCCATCTCTTTG 333
Db 440 GCACACGGTGGCCCCGATGCTGGGGAACCTCTGATCCAGCCAGCCCATCTCTTTG 499
OY 334 CGGCGGCATGATGACAGCCTTTCTGTTTCCACTTTTGCTGTACAGATGCTCTCTCC 393
Db 500 CTGGCTGATGACGACAGACCTTTCTCTGTTTGGACACAGCGAATGCTCTCTCC 559
OY 394 TGGTGGTATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 453
Db 560 TGGTGGTATGTCCTATGATCTGTACGTGGCCATCTGCCACCCCTCTCCGATATTGGCCA 619
OY 454 TCATGACCTGGAGATCTGATCAACCTCGCGGTGACTTCTTGACCACTGGAGTCTCTT 513
Db 620 TCATGACCTGGAGATCTGATCAACCTCGCGGTGACTTCTTGACCACTGGAGTCTCTT 679
OY 514 TATCTTGTATTCATCTTGTGTACTTCTTACCTTTACCTTCTGTAGGCCCA 565
Db 680 TGGCTCTGGCCCATGTGTTCTCATCTTACAGACTGCCCTTCTGTGGGCTCA 731

```

```

RESULT 11
LOCUS BE543056 580 bp mRNA linear EST 09-AUG-2000
DEFINITION 601068985F1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:3455324 5',
mRNA sequence.
ACCESSION BE543056
VERSION BE543056.1 GI:9771701
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 580)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM8441 row: 1 column: 21
High quality sequence start: 3
High quality sequence stop: 580.

```

```

FEATURES
source
1. 580
Location/Qualifiers

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3455324"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
BASE COUNT 96 a 189 c 125 g 170 t
ORIGIN

```

```

Query Match 35.4%; Score 359; DB 10; Length 580;
Best Local Similarity 80.6%; Pred. No. 3.8e-87;
Matches 444; Conservative 0; Mismatches 105; Indels 2; Gaps 2;

```

```

OY 116 GGATTCAGATGCTCCCTCTTTGGGCTCTCTCTCCCTGTCTTACGTCCTTACCCCTGCTGGGA 175
Db 1 GGATTCAGATGCTCCCTCTTTGGGCTCTCTCTCCCTGTCTTATGTCCTTACCCCTGCTGGGA 60
OY 176 ACGGACCATATCTGGGGCTCATCTCTACCTGACTCCAGACTGACGCCGCCCT-TGTACTCT 234
Db 61 ATGGACCATATCTGGGGCTCATCTCTACCTGACTCCAGACTGACGCCGCCCTATGTCCT 120
OY 235 TCCTCTACACACCTGGCGGTGCTGACATCGCCTACGCTGCAACACGGTGCCCGGATGC 294
Db 121 TCCTCTACACACCTGGCGGTGCTGACATCGCCTATGCTGCAACACAGTGCCCGGATGC 180
OY 295 TGGTGAACCTCCATGATCCAGCCAGCCCATCTCTTTGGCGGCCCATGATGACAGCCT 354
Db 181 TGGTGAACCTCCATGATCCAGCCAGCCCATCTCTTTGGCGGCCCATGATGACAGCCT 240
OY 355 TTCTGTTTTCACCTTTGCTGTACAGATGCTCTCTCTGCTGCTGCTGATGCTCTATGATC 414
Db 241 TTCTGTTTTCACCTTTGCTGTACAGATGCTCTCTCTGCTGCTGCTGATGCTCTATGATC 300
OY 415 TGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCTGCA 474
Db 301 GGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCTGCA 360
OY 475 TCACCTCGCGGTGACTTCTCTGACAGCAATGCTCTTTTATCTTATCATCTTGTGT 534
Db 361 TCACCTCGCGGTGACTTCTCTGACAGCAATGCTCTTTTATCTTATCATCTTGTGT 420
OY 535 TACTTCTACCTTATACCTTCTGTAGGCCCGAGAAATTTATCACTTTTGTGTAATC 594
Db 421 TCATCTAAGACTGCCCCCTTTGTGGGCTGTGAATATCAACACAC-TTCTTCTGTGAATC 479
OY 595 TTGGCTGTCTCAAACTTGCCCTGTGCAATATCCACATCAATGAGAATGCTTGGCC 654
Db 480 CTGTCTGTCTCAAGCCTGGCCTGTGCTGATACCTGGCTCAACGAGGTGTCATCTTGGC 539
OY 655 GGAGCAATTTTC 665
Db 540 CCTGCATGTTTC 550

```

```

RESULT 12
LOCUS AI340119/c 367 bp mRNA linear EST 29-DEC-1998
DEFINITION qk64f08.x1 NCI_CGAP_C08 Homo sapiens CDNA clone IMAGE:1873767 3',
similar to SW:OLF7_MOUSE P34984 OLFACTORY RECEPTOR-LIKE PROTEIN K7
; mRNA sequence.
ACCESSION AI340119
VERSION AI340119.1 GI:4077046
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 367)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from G1Dco
High quality sequence stop: 245.

```

```

FEATURES
source
1. 367
Location/Qualifiers

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE:1873767"
/clone_11b="NCI_CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

```

Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 1867  
High quality sequence stops: 304  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 1867 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 304.  
Location/Qualifiers  
1. .426

Query Match	35.0%	Score 355;	DB 9;	Length 367;
Best Local Similarity	100.0%	Pred. No. 3.9e-86;		
Matches 355;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 231	TTCTTTCCTTCACACACCTGGCGGTGTCGACATGCGCTACGCGCTGCAACACAGGTGCCCGG	290		
DB 367	TTCTTTCCTTCACACACCTGGCGGTGTCGACATGCGCTACGCGCTGCAACACAGGTGCCCGG	308		
QY 291	ATGCTGGTGAACCTCCTGTCATCCAGCCCAAGCCCATCTTCCTTTGCGGGCCGCATGATGCAG	350		
DB 307	ATGCTGGTGAACCTCCTGTCATCCAGCCCAAGCCCATCTTCCTTTGCGGGCCGCATGATGCAG	248		
QY 351	ACCTTTCCTGTTTCCACTTTTGCCTTCACAGATGTCTTCCTCTGCTGATGTCCTAT	410		

```

/organism="Homo sapiens"
/db_xref="GDB:3784344"
/db_xref="taxon:9606"
/clone="IMAGE:211503"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dt) primer
[5' AACTGGAAGCAATTAAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fátima Bonaudo."

```

257 ACCCTTCGCGTATTCCTCCTTTTCCTGACGAGATGATGCTCCCTCCGAGGATGATGCTCTCTT 186  
 QY 411 GATCTGTACGTGGCCATCTGGCACCCTCCGATATTGGCCATCATGACCTGGAGAGTC 470  
 |||||  
 Db 187 GATCTGTACGTGGCCATCTGGCACCCTCCGATATTGGCCATCATGACCTGGAGAGTC 128  
 QY 471 TGCATCACCCCTGGCGGTGACTCTCGAGCACACTGAGAGTCTTTTATACCTTGATTCATCTT 530  
 |||||  
 Db 127 TGCATCACCCCTGGCGGTGACTCTCGAGCACACTGAGAGTCTTTTATACCTTGATTCATCTT 68  
 QY 531 GTGTACTTCTTACCTTTACCCCTCTGTAGGCCCCAGAAATTTATCACTTTTTTT 585  
 |||||  
 Db 67 GTGTACTTCTTACCTTTACCCCTCTGTAGGCCCCAGAAATTTATCACTTTTTTT 13

Query Match	34.1%;	Score 345.6;	DB 14;	Length 426;
Best Local Similarity	96.0%;	Pred. No. 1.6e-83;		
Matches 406;	Conservative	0;	Mismatches 12;	Indels 5;
			Gaps	5;
QY 61	GACATATACATCCATCACAGAGTCTCCTACTGGGATTTCCCGTGGCCCAAGGATT	120		
DB 1	GACATATACATCCATCACAGAGTCTCCTACTGGGATTTCCCGTGGCCCAAGGATT	60		
QY 121	CAGATGCTCCTTTGGGCTCTTCGCCGTGTTCTACGTCCTTCAACCCTGTGGGAACGGG	180		
DB 61	CAGATGCTCCTTTGGGCTCTTCGCCGTGTTCTACGTCCTTCAACCCTGTGGGAACGGG	120		

RESULT 13	H62441	LOCUS	DEFINITION
	H62441	426 bp	mrna
	Y79608.r1		linear
	IMAGE:211503.5		EST 06-OCT-1995
	RECEPTOR-LIKE PROTEIN K7		mouse cdna clone
	''		olfactory
			sequence.

D<sub>b</sub>    121 ACCATACTGGGGCTCATCTCACTGAGACTCCAGAAGCGGCCCATGTACTTCTTCCTC 180  
Q<sub>y</sub>    240 TCACACCT-GGGCGTGCTGCACATCGGCTACGCTGCAACACAGGTGTCCCCGGATGCTGT 298  
       |||||    |||||    |||||    |||||    |||||    |||||    |||||    |||||    |||||    |||||  
D<sub>b</sub>    181 TCACACCTGGGGCGGTCTGCACATCGGCTTACGCTTGCAACACAGGTGTCCCCGGATGCTGT 240

ACCESSION	H62441	
VERSION	H62441.1	GI:1015273
KEYWORDS	EST.	
SOURCE	human.	

QY	299	GAACCTCCTGCATCCAGCCAGCCCATCTCTTTCGGGCGGCATGATGCAGACCTTTCT	358
Db	241	GAACCTCCTGCATCCAGCCAGCCCATCTCTTTCGGGCGCG-ATGATGCAGACCTTTCT	299

REFERENCE  
AUTHORS  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 426)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,

Db 300 GTTTCACCTTTGCTGTGCACAGAACTCTCCTCCGGGTGATGCTCTATGATCTGTA 359

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE
1 (bases 1 to 426)	Hallier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.	Generation and analysis of 280,000 human expressed sequence tags	Genome Res. 6 (9), 807-828 (1996)	97044478

OY      419 CGT-GGCCATCTGCCACCCCCCTCGGATATT-  
          ||| ||||||| ||||||||||| |||||||| | ||| |  
Db      360 CGTGGCCATCTGCNACCCCCCTCGGATATTNGGCCATCATGAACNGGGAGAGTCTGNC  
  
OY      477 ACC 479  
          | |  
Db      420 ATC 422

**TITLE** Generation and analysis of 280,000 human expressed sequence tags  
**JOURNAL** Genome Res. 6 (9), 807-828 (1996)  
**MEDLINE** 97044478  
**COMMENT** Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800

RESULT 1A				
BM670360				
LOCUS	623 bp	mrna	linear	EST 27-FEB-2002
DEFINITION	BM670360			
	UI-E-DWI-ahc-b-11-0-UI.s1	UI-E-DWI	homo sapiens	cdna clone







BASE COUNT 115 a 73 c 110 g 63 t  
ORIGIN

Query Match 33.3%; Score 337.4; DB 9; Length 361;  
Best Local Similarity 96.9%; Pred. No. 2.5e-81;  
Matches 344; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

QY 231 TTTCTTCCTCTCACACCTGGGGGTGTCGACATCGCTACGCTGCAACAGGTGCCCGG 290
    ||| ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 TTTTCCCTCTCACACCTGGGGGTGTCGACATGCTACGCTGCAACAGGTGCCCGG 302

QY 291 ATGCTGTGAACCTCCGTCATCCAGCCAGCCCATCTCTTGGCGGGCGCATGATGAG 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 ATGCTGTGAACCTCCGTCATCCAGCCAGCCCATCTCTTGGCGGGCGCATGATGAG 242

QY 351 ACCTTTCTGTTTTCACATTTGCTGTACAGAAATGCTCTCTGTTGATGATGCTAT 410
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 ACCTTTCTGTTTTCACATTTGCTGTACAGAAATGCTCTCTGTTGATGATGCTAT 182

QY 411 GATCTGTAGGTGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGAGAGTC 470
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GATTTGTAGGTGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGAGAGTC 122

QY 471 TGCATCACCCCTCGCGGTGACTTCTGAGACCACTGAGTCTTTATCCTGATTCATCTT 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TGCATCACCCCTCGCGGTGACTTCTGAGACCACTGAGTCTTTATCCTGATTCATCTT 62

QY 531 GTGTACTTCTACCTTTACCTTCTGTAGGCCCCAGAAATTATCACTTTTTP 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GTGTACTTCTACCTTTACCTTCTGTAGGCCCCAGAAATTATCACTTTTTP 7
    
```

Search completed: May 2, 2003, 06:14:26  
Job time : 1553 secs

•  
•  
•  
•

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 04:33:51 ; Search time 2809 Seconds  
(without alignments)  
10505.605 Million cell updates/sec

Title: US-09-898-586-23

Perfect score: 1014  
Sequence: 1 taacacattctcctaacc.....tatggcattgtgactgacaa 1014

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl : \*  
1: gb\_ba : \*  
2: gb\_htg : \*  
3: gb\_in : \*  
4: gb\_cm : \*  
5: gb\_ov : \*  
6: gb\_pat : \*  
7: gb\_ph : \*  
8: gb\_pl : \*  
9: gb\_pr : \*  
10: gb\_ro : \*  
11: gb\_sts : \*  
12: gb\_sy : \*  
13: gb\_un : \*  
14: gb\_vl : \*  
15: em\_ba : \*  
16: em\_fun : \*  
17: em\_hum : \*  
18: em\_in : \*  
19: em\_mu : \*  
20: em\_cm : \*  
21: em\_or : \*  
22: em\_ov : \*  
23: em\_pat : \*  
24: em\_ph : \*  
25: em\_pl : \*  
26: em\_ro : \*  
27: em\_sts : \*  
28: em\_un : \*  
29: em\_vl : \*  
30: em\_htg\_hum : \*  
31: em\_htg\_inv : \*  
32: em\_htg\_other : \*  
33: em\_htg\_mus : \*  
34: em\_htg\_pln : \*  
35: em\_htg\_rod : \*  
36: em\_htg\_mam : \*  
37: em\_htg\_vrt : \*  
38: em\_sy : \*  
39: em\_htgo\_hum : \*  
40: em\_htgo\_mus : \*  
41: em\_htgo\_other : \*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1014	100.0	1014	6	AX197458	AX197458 Sequence
2	990	97.6	130030	9	AC004889	AC004889 Homo sapi
3	987.4	97.4	1012	6	AX197456	AX197456 Sequence
4	983.6	97.0	1040	6	AX197438	AX197438 Sequence
5	983.6	97.0	1040	6	AX197469	AX197469 Sequence
6	983.6	97.0	59130	9	AL135904	AL135904 Human DNA
7	983.6	97.0	140915	9	AC005587	AC005587 Homo sapi
8	936	92.3	974	6	AX078373	AX078373 Sequence
9	905.4	89.3	930	6	AX242191	AX242191 Sequence
10	903.6	89.1	933	6	AX451262	AX451262 Sequence
11	900.6	88.8	930	6	AX241442	AX241442 Sequence
12	900.6	88.8	930	6	AX241473	AX241473 Sequence
13	675	66.6	104955	9	AC091768	AC091768 Homo sapi
14	645.4	63.6	204316	2	AC091746	AC091746 Mus muscu
15	637.4	62.9	650	9	AF399598	AF399598 Homo sapi
16	622	61.3	933	10	AY073444	AY073444 Mus muscu
17	573.2	56.5	933	10	AY073100	AY073100 Mus muscu
18	549.8	54.2	1262	6	AX468421	AX468421 Sequence
19	546.6	53.9	66610	9	AC074386	AC074386 Homo sapi
20	545	53.7	1315	9	AB065692	AB065692 Homo sapi
21	535.4	52.8	930	6	AX242188	AX242188 Sequence
22	519.8	51.3	813	6	AX451240	AX451240 Sequence
23	518.4	51.1	1792	6	AX078369	AX078369 Sequence
24	517.8	51.1	993	6	AX242190	AX242190 Sequence
25	509.8	50.3	542	6	AX451241	AX451241 Sequence
26	508.8	50.2	542	6	AX451242	AX451242 Sequence
27	481.2	47.5	1713	6	AR092423	AR092423 Sequence
28	458	45.2	106981	2	AC120563	AC120563 Rattus no
29	444.4	43.8	162320	2	AC076959	AC076959 Homo sapi
30	442.2	43.6	667	10	MUSODORECA	L14566 Mouse (K7)
31	441.8	43.6	966	10	AY073724	AY073724 Mus muscu
32	440.2	43.4	966	10	AY073723	AY073723 Mus muscu
33	440.2	43.4	147341	2	AC106354	AC106354 Rattus no
34	437.8	43.2	10985	2	AC090838	AC090838 Homo sapi
35	430	42.4	933	10	AY073820	AY073820 Mus muscu
36	429.6	42.4	576	6	AX242035	AX242035 Sequence
37	426.8	42.1	1357	9	AB065695	AB065695 Homo sapi
38	423.8	41.8	942	10	AY073443	AY073443 Mus muscu
39	423.6	41.8	162320	2	AC076959	AC076959 Homo sapi
40	410.4	40.5	424	6	AX451249	AX451249 Sequence
41	409.6	40.4	1334	9	AB065480	AB065480 Homo sapi
42	409.6	40.4	130030	9	AC004889	AC004889 Homo sapi
43	409.2	40.4	942	6	AX241995	AX241995 Sequence
44	409.2	40.4	1207	9	AB065691	AB065691 Homo sapi
45	407.6	40.2	1333	9	AB065693	AB065693 Homo sapi

#### ALIGNMENTS

RESULT 1  
AX197458  
LOCUS AX197458 1014 bp DNA  
DEFINITION Sequence 23 from Patent WO0151632.  
ACCESSION AX197458  
VERSION AX197458.1 GI:15387838  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1014)  
Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T.,  
Spytek,K.A. and Li,L.  
TITLE  
Odorant receptor polypeptides and nucleic acids encoding same

JOURNAL Patent: WO 0151632-A 23 19-JUL-2001;

Curagen Corporation (US)

FEATURES Location/Qualifiers

1. 1014

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 218 a 281 c 216 g 299 t

ORIGIN

Query Match 100.0%; Score 1014; DB 6; Length 1014;

Best Local Similarity 100.0%; Pred. No. 3.2e-239;

Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAACACTTCTCCAAACCATGAGCATTAATGATTTCCCTCTGTCATAGGATATGGG 60  
Db 1 TAAACACTTCTCTAAACCATGAGCATTAATGATTTCCCTCTGTCATAGGATATGGG 60  
QY 61 GACAATATACATCAATCAGAGAGTCCCTCTACTGGATTCCCTGGCCCAAGAT 120  
Db 61 GACAATATACATCAATCAGAGAGTCCCTCTACTGGATTCCCTGGCCCAAGAT 120  
QY 121 CAGATGCTCTCTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
Db 121 CAGATGCTCTCTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
QY 181 ACCATAGTGGGGCTATCTCACTGAGCTCAGACTGACGCCCCCTGACTCTCTCT 240  
Db 181 ACCATAGTGGGGCTATCTCACTGAGCTCAGACTGACGCCCCCTGACTCTCTCT 240  
QY 241 CACACCTGGGGCTCTGACATGCGCTAGCGCTGCAACACGCGTGGCCCGGATG 300  
Db 241 CACACCTGGGGCTCTGACATGCGCTAGCGCTGCAACACGCGTGGCCCGGATG 300  
QY 301 ACCTCTGATCCAGCCAGCCCATCTCTTGGCGGCGCATGATGACACCTTTCT 360  
Db 301 ACCTCTGATCCAGCCAGCCCATCTCTTGGCGGCGCATGATGACACCTTTCT 360  
QY 361 TTTCACCTTTGGCTGACAGATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
Db 361 TTTCACCTTTGGCTGACAGATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
QY 421 TGGCCATCTGACACCCCTCCGATTTGGCCATCATGACCTGAGAGTCTGCATCAC 480  
Db 421 TGGCCATCTGACACCCCTCCGATTTGGCCATCATGACCTGAGAGTCTGCATCAC 480  
QY 481 TCGGGGTGACTCTCTGAGACCACTGAGTCTTTATCTTATCATCTTGTACTTC 540  
Db 481 TCGGGGTGACTCTCTGAGACCACTGAGTCTTTATCTTATCATCTTGTACTTC 540  
QY 541 TACCTTTACCTCTCTGAGGCCCCAGAAATTTATCACTTTTGTGAAATCTGGCT 600  
Db 541 TACCTTTACCTCTCTGAGGCCCCAGAAATTTATCACTTTTGTGAAATCTGGCT 600  
QY 601 GTTCTCAACTTGGCTGTGAGATACCCACATCAATGAGACATGCTTGGCCGAGCA 660  
Db 601 GTTCTCAACTTGGCTGTGAGATACCCACATCAATGAGACATGCTTGGCCGAGCA 660  
QY 661 ATTCTGGGCTGTGGGACCCCTGTCCACAATTTAGTTTCAATATGTCATCTCTGT 720  
Db 661 ATTCTGGGCTGTGGGACCCCTGTCCACAATTTAGTTTCAATATGTCATCTCTGT 720  
QY 721 GCTATCTTCAAGATCAATCAAGGAGTTCAGAGAAAGCCCTTGCACCTCTCTCC 780  
Db 721 GCTATCTTCAAGATCAATCAAGGAGTTCAGAGAAAGCCCTTGCACCTCTCTCC 780  
QY 781 CACCTCTGTGATGAGCTCTTTATGGACAGCCATATCATATATGTTGACCCAGA 840  
Db 781 CACCTCTGTGATGAGCTCTTTATGGACAGCCATATCATATATGTTGACCCAGA 840  
QY 841 TATGGGAACCCCAAGAGCAGAGAAATATCTCTCTCTCTCTCTCTCTCTCTCT 900  
Db 841 TATGGGAACCCCAAGAGCAGAGAAATATCTCTCTCTCTCTCTCTCTCTCTCT 900

QY 901 ATGCTCAATCCCTTATCTGTAGTCTTAGGAAGTCAAGATGAGATCTTTGAGAGA 960  
Db 901 ATGCTCAATCCCTTATCTGTAGTCTTAGGAAGTCAAGATGAGATCTTTGAGAGA 960  
QY 961 GTGCTGGAGTAGAAGGGCTTTAAGAAAGATATGCGATTTGACTGACAGAA 1014  
Db 961 GTGCTGGAGTAGAAGGGCTTTAAGAAAGATATGCGATTTGACTGACAGAA 1014

RESULT 2

AC004889/c

LOCUS AC004889 130030 bp DNA linear PRI 14-JAN-1999

DEFINITION Homo sapiens clone DJ0798C17, complete sequence.

AC004889

AC004889.1 GI:4156187

HTG.

KEYWORDS Homo sapiens.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 130030)

Waterston,R.H.  
The sequence of Homo sapiens clone

unpublished  
2 (bases 1 to 130030)

Waterston,R.H.  
Direct Submission

Submitted (12-JUN-1998) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA

REFERENCE 3 (bases 1 to 130030)

Waterston,R.H.  
Direct Submission

Submitted (14-JAN-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA

On Jan 14, 1999 this sequence version replaced gi:3213108.

COMMENT

FEATURES

source

1. 130030

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="DJ0798C17"

BASE COUNT 38117 a 25410 c 26567 g 39936 t

ORIGIN

Query Match

Best Local Similarity 97.6%; Score 990; DB 9; Length 130030;

Matches 1012; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 TAAACACTTCTCTAAACCATGAGCATTAATGATTTCCCTCTGTCATAGGATATGGG 60  
Db 111297 TAAACACTTCTCTAAACCATGAGCATTAATGATTTCCCTCTGTCATAGGATATGGG 111238  
QY 61 GACAATATACATCCATCAGAGAGTTCCTCTACTGGGATTTCCCGTGGCCAGAGAT 120  
Db 111237 GACAATATACATCCATCAGAGAGTTCCTCTACTGGGATTTCCCGTGGCCAGAGAT 111178  
QY 121 CAGATGCTCTCTTGGGCTCTTCCCTGTCTAGCTTCAACCCCTGCTGGGGAACGG 180  
Db 111177 CAGATGCTCTCTTGGGCTCTTCCCTGTCTAGCTTCAACCCCTGCTGGGGAACGG 111118  
QY 181 ACCATAGTGGGGCTCATCTCACTGAGCTCAGAGTGCAGAGCCCCC-TGTACTTCTCTC 239  
Db 111117 ACCATAGTGGGGCTCATCTCACTGAGCTCAGAGTGCAGAGCCCCC-CATGTACTTCTCTC 111058  
QY 240 TCACACCTGGGGTGTGTCAGATGCGCTACGCGCTGCAACAGGGTGCCCGGATGCTGTG 299  
Db 111057 TCACACCTGGGGTGTGTCAGATGCGCTACGCGCTGCAACAGGGTGCCCGGATGCTGTG 110998  
QY 300 AACCTCTGATCCAGCCAGCCCATCTCTTGGCGGCGCATGATGACAGACTTTCTG 359  
Db 110997 AACCTCTGATCCAGCCAGCCCATCTCTTGGCGGCGCATGATGACAGACTTTCTG 110938

QY 360 TTTTCACTTTTGTGTGACAGAAATGTCCTCTGCTGTGATGTCCTAAGATCTGTAC 419  
|||||  
Db 110937 TTTTCACTTTTGTGTGACAGAAATGTCCTCTGCTGTGATGTCCTAAGATCTGTAC 110878  
QY 420 GTGGCATTGTGACACCCCTCCGATATTGGCCATCATGACCTGGAGAGCTGCATCACC 479  
|||||  
Db 110877 GTGGCATTGTGACACCCCTCCGATATTGGCCATCATGACCTGGAGAGCTGCATCACC 110818  
QY 480 CTGGGCTGACTTCTCTGACCACTGGAGTCTTTATCTTATCTGATTCATCTGTGTACTT 539  
|||||  
Db 110817 CTGGGCTGACTTCTCTGACCACTGGAGTCTTTATCTTATCTGATTCATCTGTGTACTT 110758  
QY 540 CTACCTTTACCTCTCTGTAGGCCCCAGAAAATTTATCACTTTTGTGAAATCTTGGC 599  
|||||  
Db 110757 CTACCTTTACCTCTCTGTAGGCCCCAGAAAATTTATCACTTTTGTGAAATCTTGGC 110699  
QY 600 TGTCTCAAACTTGCCCTGTGAGATACCCACATCAATGAGACATGGTCTGGCCGAGC 659  
|||||  
Db 110698 TGTCTCAAACTTGCCCTGTGAGATACCCACATCAATGAGACATGGTCTGGCCGAGC 110639  
QY 660 AATTCTGGGCTGTGTGGACCCCTGTCCACAATGTAGTTTCATATATGTCATCTCTG 719  
|||||  
Db 110638 AATTCTGGGCTGTGTGGACCCCTGTCCACAATGTAGTTTCATATATGTCATCTCTG 110579  
QY 720 TGTCTCACTTCAGATCCAAATCAAGGGAAGTTCAGAGGAAAGCCCTCTGACCTGCTCTC 779  
|||||  
Db 110578 TGTCTCACTTCAGATCCAAATCAAGGGAAGTTCAGAGGAAAGCCCTCTGACCTGCTCTC 110519  
QY 780 CCACCTCTGTGTGATGTGACTCTTTATGGCAGACCATTATCATGTATGTGAGCCAG 839  
|||||  
Db 110518 CCACCTCTGTGTGATGTGACTCTTTATGGCAGACCATTATCATGTATGTGAGCCAG 110459  
QY 840 ATATGGGAACCCCAAGAGAGAGAAATATCTCCTGCTGTTTCACAGCCCTTTAATCC 899  
|||||  
Db 110458 ATATGGGAACCCCAAGAGAGAGAAATATCTCCTGCTGTTTCACAGCCCTTTAATCC 110399  
QY 900 CATGCTCAATCCCTTATCTGTACTCTTAGGACTCAGAAGTGAAGAATACTTTGAAGAG 959  
|||||  
Db 110398 CATGCTCAATCCCTTATCTGTACTCTTAGGACTCAGAAGTGAAGAATACTTTGAAGAG 110339  
QY 960 AGTGTGGGAGTAGAAGGGCTTTATGAAAAGGATTTATGGCATTTGTACTGACA 1013  
|||||  
Db 110338 AGTGTGGGAGTAGAAGGGCTTTATGAAAAGGATTTATGGCATTTGTACTGACA 110285

RESULT 3  
AX197456 1012 bp DNA linear PAT 29-AUG-2001  
LOCUS AX197456  
DEFINITION Sequence 21 from Patent WO0151632.  
ACCESSION AX197456  
VERSION AX197456.1 GI:15387837  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1012)  
AUTHORS Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T.,  
TITLE Odorant receptor polypeptides and nucleic acids encoding same  
JOURNAL Patent: WO 0151632-A 21 19-JUL-2001;  
FEATURES  
source location/Qualifiers  
1..1012  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 219 a 281 c 215 g 297 t  
ORIGIN

Query Match 97.4%; Score 987.4; DB 6; Length 1012;  
Best Local Similarity 99.7%; Pred. No. 1.2e-232;  
Matches 1010; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 2 AAACACTTCTCCAAACCATGAGCATTAACCTTGATTTCTCTGTGCATAGGATATGGGG 61  
|||||  
Db 1 AAACACTTCTCCAAACCATGAGCATTAACCTTGATTTCTCTGTGCATAGGATATGGAG 60  
QY 62 ACAATATACATGCATCACAGAGTTCCTCTACTGGGATTTCCCGTTGGCCCAAGATTC 121  
|||||  
Db 61 ACAATATACATGCATCACAGAGTTCCTCTACTGGGATTTCCCGTTGGCCCAAGATTC 120  
QY 122 AGATGCTCTCTTTGGGCTCTTCTCCCTGTCTACGTCTTACCCCTGTGGGAACGGGA 181  
|||||  
Db 121 AGATGCTCTCTTTGGGCTCTTCTCCCTGTCTACGTCTTACCCCTGTGGGAACGGGA 180  
QY 182 CCATACCTGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCC-TGTACTTCTCTCT 240  
|||||  
Db 181 CCATACCTGGGCTCATCTCACTGGACTCCAGACTGCACGCCCCCAGTACTTCTCTCT 240  
QY 241 CACACCTGGGCTGTGACATCGCCCTACGCTGCAACACGGTGGCCCGATGCTGTGA 300  
|||||  
Db 241 CACACCTGGGCTGTGACATCGCCCTACGCTGCAACACGGTGGCCCGATGCTGTGA 300  
QY 301 ACCTCTGCATCCAGCCCAAGCCCATCTCTTTGGCGGGCCGATGATGACACCTTTCTGT 360  
|||||  
Db 301 ACCTCTGCATCCAGCCCAAGCCCATCTCTTTGGCGGGCCGATGATGACACCTTTCTGT 360  
QY 361 TTTCCACTTTTGTCTGTACAGAAATGTCTCTCTCTGCTGTGTGATGTCTATGATCTGTACG 420  
|||||  
Db 361 TTTCCACTTTTGTCTGTACAGAAATGTCTCTCTCTGCTGTGTGATGTCTATGATCTGTACG 420  
QY 421 TGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480  
|||||  
Db 421 TGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCTGCATCACCC 480  
QY 481 TGGGGGTGACTTCCCTGAGACCCACTGGAGTCTTTTATCTCTGTATTCATCTGTGTACTTC 540  
|||||  
Db 481 TGGGGGTGACTTCCCTGAGACCCACTGGAGTCTTTTATCTCTGTATTCATCTGTGTACTTC 540  
QY 541 TACCTTTACCTCTCTGTAGGCCCCCAGAAAATTTATCACCTTTTGTGAAATCTTGGCT 600  
|||||  
Db 541 TACCTTTACCTCTCTGTAGGCCCCCAGAAAATTTATCACCTTTTGTGAAATCTTGGCT 599  
QY 601 GTTCTCAAACTTGCCCTGTGACAGATACCCACATCAATGAGAACATGGTCTTGGCCGAGCA 660  
|||||  
Db 600 GTTCTCAAACTTGCCCTGTGACAGATACCCACATCAATGAGAACATGGTCTTGGCCGAGCA 659  
QY 661 ATTTCTGGGCTGTGGGACCCCTGTGCCACAAATTTAGTTTCANATATGTGCATCTCTGT 720  
|||||  
Db 660 ATTTCTGGGCTGTGGGACCCCTGTGCCACAAATTTAGTTTCANATATGTGCATCTCTGT 719  
QY 721 GCTATCTCTCAGATCAATCAAGGGAAGTTCAGAGGAAAGCCCTGTGCACCTGTCTCC 780  
|||||  
Db 720 GCTATCTCTCAGATCAATCAAGGGAAGTTCAGAGGAAAGCCCTGTGCACCTGTCTCC 779  
QY 781 CACCTCTGTGTGATGTGACTCTTTATGCGACAGCCATTTATCATGTATGTGAGCCAGGA 840  
|||||  
Db 780 CACCTCTGTGTGATGTGACTCTTTATGCGACAGCCATTTATCATGTATGTGAGCCAGGA 839  
QY 841 TATGGGAACCCCAAGAGAGAGAAATATCTCTCTGTGTTTCAGAGCCCTTTAATGCC 900  
|||||  
Db 840 TATGGGAACCCCAAGAGAGAGAAATATCTCTCTGTGTTTCAGAGCCCTTTAATGCC 899  
QY 901 ATGCTCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAAGTGAAGAATACTTTGAAGAGA 960  
|||||  
Db 900 ATGCTCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAAGTGAAGAATACTTTGAAGAGA 959  
QY 961 GTGCTGGGAGTAGAAGGGCTTTATGAAAAGGATTTATGGCATTTGTACTGACAGCA 1013  
|||||  
Db 960 GTGCTGGGAGTAGAAGGGCTTTATGAAAAGGATTTATGGCATTTGTACTGACAGCA 1012

RESULT 4  
AX197438 1040 bp DNA linear PAT 29-AUG-2001  
LOCUS AX197438  
DEFINITION Sequence 3 from Patent WO0151632.



ACCESSION AX197438  
VERSION AX197438.1 GI:15387828  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1040)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE Odorant receptor polypeptides and nucleic acids encoding same  
JOURNAL Patent: WO 0151632-A 3 19-JUL-2001;  
FEATURES  
source 1.1040  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 229 a 285 c 221 g 305 t  
ORIGIN

Query Match 97.0%; Score 983.6; DB 6; Length 1040;  
Best Local Similarity 99.4%; Pred. No. 1e-231;  
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

OY 1 TAAACACTTCTCCTAAACATGAGCATTAACCTGATTCCTCTGTCATAGGATATGGG 60  
DB 28 TAAACACTTCTCCTAAACATGAGCATTAACCTGATTCCTCTGTCATAGGATATGGG 87  
OY 61 GACAATATTAACATCCATCAAGAGTTCCTCTCTACTGGGATTTCCCGTGGCCCAAGATT 120  
DB 88 GACAATATTAACATCCATCAAGAGTTCCTCTCTACTGGGATTTCCCGTGGCCCAAGATT 147  
OY 121 CAGATGCTCCTCTTTGGGCTCTCTCCCTGTTCTAGCTCTTCACCCCTGCTGGGAACGG 180  
DB 148 CAGATGCTCCTCTTTGGGCTCTCTCCCTGTTCTAGCTCTTCACCCCTGCTGGGAACGG 207  
OY 181 ACCATACTGGGGCTCATCTGACTGGACTCCAGATGCCAGCCGCC-1GTACTTCTCTCTC 239  
DB 208 ACCATACTGGGGCTCATCTGACTGGACTCCAGATGCCAGCCGCCATGACTTCTCTCTC 267  
OY 240 TCACACCTGGGGCTGCTGACATGCGCTACGCTGCAACACGCGTGGCCCGGATGCTGCTG 299  
DB 268 TCACACCTGGGGCTGCTGACATGCGCTACGCTGCAACACGCGTGGCCCGGATGCTGCTG 327  
OY 300 AACCTCCTGCATCCAGCCAGCCCATCTCTCTTTGGGGCCCGCATGATGAGACCTTTCTG 359  
DB 328 AACCTCCTGCATCCAGCCAGCCCATCTCTCTTTGGGGCCCGCATGATGAGACCTTTCTG 387  
OY 360 TTTTCCACTTTTGTGCTGACAGAAATGTCTCTCTCTGCTGATGCTCTATGATCTGTAC 419  
DB 388 TTTTCCACTTTTGTGCTGACAGAAATGTCTCTCTCTGCTGATGCTCTATGATCTGTAC 447  
OY 420 GTGGCCATCTGACACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCAC 479  
DB 448 GTGGCCATCTGACACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCAC 507  
OY 480 CTGCGGGTGAATCTCTGAGACCTGAGAGTCTTTATCTTATCTGATTCATCTGTCTACT 539  
DB 508 CTGCGGGTGAATCTCTGAGACCTGAGAGTCTTTATCTTATCTGATTCATCTGTCTACT 567  
OY 540 CTACCTTTACCTCTCTGAGAGCCCGAGAAATTTATCACTTTTCTTTGTAATCTTGGC 599  
DB 568 CTACCTTTACCTCTCTGAGAGCCCGAGAAATTTATCAC-TTTTCTTTGTAATCTTGGC 626  
OY 600 TGTTCACAACCTTGCTGAGATACCCATCATATGAGACATGCTGCTGGCCGAGC 659  
DB 627 TGTTCACAACCTTGCTGAGATACCCATCATATGAGACATGCTGCTGGCCGAGC 686  
OY 660 AATTTCTGGGCTGCTGGAGCCTTGTCCACAATTTAGTTTCATATATGATGCTCTCTG 719  
DB 687 AATTTCTGGGCTGCTGGAGCCTTGTCCACAATTTAGTTTCATATATGATGCTCTCTG 746  
OY 720 TGTTCACCTTTGCTGCTGACAGAAATGTCTCTCTCTGCTGATGCTCTGCTCTCTC 779

|||||  
DB 747 TGCTATCCTTCAGATCCATCAAGGAGTTCAGAGAAAGCTTCCGACCTGCTCTC 806  
OY 780 CCACCTCTGTGTGATTTGAGACTCTTTATGCGACAGCCATATCATGTATGTGACCCAG 839  
DB 807 CCACCTCTGTGTGATTTGAGACTCTTTATGCGACAGCCATATCATGTATGTGACCCAG 866  
OY 840 ATATGGGAACCCCAAGAGCAGAGAATATCTCTGCTGTTTCACAGCCTCTTTAATCC 899  
DB 867 ATATGGGAACCCCAAGAGCAGAGAATATCTCTGCTGTTTCACAGCCTCTTTAATCC 926  
OY 900 CATGCTCAATCCCTTATCTGTAGTCTTGAAGACTCAGAAAGTGAAGATCTTTGAAGAG 959  
DB 927 CATGCTCAATCCCTTATCTGTAGTCTTGAAGACTCAGAAAGTGAAGATCTTTGAAGAG 986  
OY 960 AGTGCTGGAGTAGAAGAGGCTTTATGAAGGATTTATGCGATTTGACTGACA 1013  
DB 987 AGTGCTGGAGTAGAAGAGGCTTTATGAAGGATTTATGCGATTTGACTGACA 1040

RESULT 5  
AX197469  
LOCUS AX197469 1040 bp DNA linear PAT 29-AUG-2001  
DEFINITION Sequence 34 from Patent WO0151632.  
ACCESSION AX197469  
VERSION AX197469.1 GI:15387841  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1040)  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE Odorant receptor polypeptides and nucleic acids encoding same  
JOURNAL Patent: WO 0151632-A 34 19-JUL-2001;  
FEATURES  
source 1.1040  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 229 a 285 c 221 g 305 t  
ORIGIN

Query Match 97.0%; Score 983.6; DB 6; Length 1040;  
Best Local Similarity 99.4%; Pred. No. 1e-231;  
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

OY 1 TAAACACTTCTCCTAAACATGAGCATTAACCTGATTCCTCTGTCATAGGATATGGG 60  
DB 28 TAAACACTTCTCCTAAACATGAGCATTAACCTGATTCCTCTGTCATAGGATATGGG 87  
OY 61 GACAATATTAACATCCATCAAGAGTTCCTCTCTACTGGGATTTCCCGTGGCCCAAGATT 120  
DB 88 GACAATATTAACATCCATCAAGAGTTCCTCTCTACTGGGATTTCCCGTGGCCCAAGATT 147  
OY 121 CAGATGCTCCTCTTTGGGCTCTCTCCCTGTTCTAGCTCTTCACCCCTGCTGGGAACGG 180  
DB 148 CAGATGCTCCTCTTTGGGCTCTCTCCCTGTTCTAGCTCTTCACCCCTGCTGGGAACGG 207  
OY 181 ACCATACTGGGGCTCATCTGACTGGACTCCAGACTGCAGCCGCC-1GTACTTCTCTCTC 239  
DB 208 ACCATACTGGGGCTCATCTGACTGGACTCCAGACTGCAGCCGCCCATGACTTCTCTCTC 267  
OY 240 TCACACCTGGGGCTGCTGACATGCGCTACGCTGCAACACGCGTGGCCCGGATGCTGCTG 299  
DB 268 TCACACCTGGGGCTGCTGACATGCGCTACGCTGCAACACGCGTGGCCCGGATGCTGCTG 327  
OY 300 AACCTCCTGCATCCAGCCAGCCCATCTCTCTTTGGGGCCCGCATGATGAGACCTTTCTG 359  
DB 328 AACCTCCTGCATCCAGCCAGCCCATCTCTCTTTGGGGCCCGCATGATGAGACCTTTCTG 387  
OY 360 TTTTCCACTTTTGTGCTGACAGAAATGTCTCTCTCTGCTGATGCTCTATGATCTGTAC 419

```

Db      388  TTTTCACCTTTTGTCTGTACAGAAATGTCCTCTGGTGATGATGATCTGTAC 447
Qy      420  GTGCCATCTGCCACCCCTCCGATATTTGGCATCATGACCTGAGAGTCTGCATCAC 479
Db      448  GTGGCATCTGCCACCCCTCCGATATTTGGCATCATGACCTGAGAGTCTGCATCAC 507
Qy      480  CTGGGCTGACTTCTGTGACCACTGGAGTCTTTATCTTATCTTATCTTATCTTACTT 539
Db      508  CTCGGGAGACTTCTGTGACCACTGGAGTCTTTATCTTATCTTATCTTATCTTACTT 567
Qy      540  CTACCTTACCTTCTGTGACCCCGAGAAATTTATCACTTTTGTGTAATCTTGGC 599
Db      568  CTACCTTACCTTCTGTGACCCCGAGAAATTTATCAC-TTTTGTGTAATCTTGGC 626
Qy      600  TGTCTCAAACTTGCCTGTGACATACCCACATCAATGAGACATGGTCTGGCCGAGC 659
Db      627  TGTCTCAAACTTGCCTGTGACATACCCACATCAATGAGACATGGTCTGGCCGAGC 686
Qy      660  AATTCTGGGCTGTGGACCTTGTCCACAATTTAGTTTCATATATGATCTCTG 719
Db      687  AATTCTGGGCTGTGGACCTTGTCCACAATTTAGTTTCATATATGATCTCTG 746
Qy      720  TGTATCTTCAATCAAGGAGAGTTCAGAGAAAGCCTTCTGACCTCTCTC 779
Db      747  TGTATCTTCAATCAAGGAGAGTTCAGAGAAAGCCTTCTGACCTCTCTC 806
Qy      780  CCACCTCTGTGTATGTGACTTTTATGACACAGCATTTATCATATGATCTTAC 839
Db      807  CCACCTCTGTGTATGTGACTTTTATGACACAGCATTTATCATATGATCTTAC 866
Qy      840  ATATGGGAACCCCAAGAGAGAAATATATCTCTGCTTTCACAGCCTTTAATCC 899
Db      867  ATATGGGAACCCCAAGAGAGAAATATATCTCTGCTTTCACAGCCTTTAATCC 926
Qy      900  CATGCTCAATCCCTTATCTGTAGTCTTAGAAGTCAAGAGTGAAGATCTTTGAAG 959
Db      927  CATGCTCAATCCCTTATCTGTAGTCTTAGAAGTCAAGAGTGAAGATCTTTGAAG 986
Qy      960  AGTGTGGAGTAGAAGGCTTTATGAAAGGATTTAGCATTTGTGACTGACA 1013
Db      987  AGTGTGGAGTAGAAGGCTTTATGAAAGGATTTAGCATTTGTGACTGACA 1040

```

```

RESULT 6
AL135904/c 59130 bp DNA linear PRI 20-JUL-2000
LOCUS      Human DNA sequence from clone Rps-1005H11 on chromosome 6 contains
DEFINITION part of the gene for a 7 transmembrane receptor (rhodopsin family)
              (olfactory receptor like) protein, the gene for WUGSC:H_DJ0988615.3
              protein, part of the PDNP3 (phosphodiesterase 1/nucleotide
              pyrophosphatase 3) gene, ESTs, STSS and GSSs, complete sequence.
ACCESSION  AL135904
VERSION     AL135904.11 GI:7159399
KEYWORDS   HTG.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 59130)
AUTHORS     Lovell, J.
TITLE       Direct Submission
JOURNAL     Submitted (05-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone
            requests: clonerequests@sanger.ac.uk
COMMENT     On Mar 6, 2000 this sequence version replaced gi:7105768.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria

```

as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/chr6>

RP5-1005H11 is from the library RPCI-5 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR:pcypac2>

## FEATURES

source

```

repeat_region
  /note="Alus repeat: matches 1. .68 of consensus"
  /location/Qualifiers
    1. .59130
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="6"
    /clone="RP5-1005H11"
    /clone_1lb="RPCI-5"
    1. .68
    /note="Alus repeat: matches 1. .68 of consensus"
    546. .844
    /note="AluJo repeat: matches 2. .301 of consensus"
    870. .940
    /note="U4 repeat: matches 1. .72 of consensus"
    1691. .2906
    /note="L2 repeat: matches 1425. .2737 of consensus"
    3457. .3960
    /note="HAL1 repeat: matches 420. .971 of consensus"
    3971. .4106
    /note="MLR1D repeat: matches 1. .129 of consensus"
    4249. .4282
    /note="17 copies 2 mer tg 100% conserved"
    4966. .5147
    /note="AluJb repeat: matches 1. .182 of consensus"
    5216. .5307
    /note="L2 repeat: matches 2405. .2503 of consensus"
    5495. .5794
    /note="Alusx repeat: matches 1. .296 of consensus"
    5891. .6180
    /note="Alusx repeat: matches 1. .291 of consensus"
    7607. .7618
    /note="L1M1 repeat: matches 4856. .4867 of consensus"
    7619. .8347
    /note="L1PA16 repeat: matches 5421. .6156 of consensus"
    8348. .8898
    /note="L1M1 repeat: matches 4866. .5403 of consensus"
    9018. .9866
    /note="TIGGER1 repeat: matches 1. .836 of consensus"
    9867. .10189
    /note="Alusx repeat: matches 1. .309 of consensus"
    10190. .11723
    /note="TIGGER1 repeat: matches 836. .2418 of consensus"
    12404. .12715
    /note="Alusx repeat: matches 1. .312 of consensus"
    complement(13947. .14176)
    /note="match: GSS: Em:B74942"
    14280. .14383
    /note="Alusg/x repeat: matches 190. .294 of consensus"
    15340. .15508
    /note="L2 repeat: matches 2563. .2750 of consensus"

```

```
repeat_region 15684..15968
/note="Alusx repeat: matches 35. .312 of consensus"
repeat_region 16527..16612
/note="MIR repeat: matches 50. .138 of consensus"
repeat_region 16731..17027
/note="Alusg repeat: matches 1. .295 of consensus"
repeat_region 17609..17747
/note="MER91A repeat: matches 1. .153 of consensus"
repeat_region 17867..18162
/note="Aluy repeat: matches 1. .299 of consensus"
repeat_region 18461..18950
/note="MER1A repeat: matches 2. .527 of consensus"
repeat_region 19185..19224
/note="20 copies 2 mer tg 85% conserved"
misc_feature 19474..19712
/note="match: STS: Em:G16149"
repeat_region 20037..20361
/note="Alusx repeat: matches 1. .311 of consensus"
gene complement(21566..22384)
/gene="dJ1005H11.1"
complement(21566..>22384)
/gene="dJ1005H11.1"
CDS /note="match: proteins: Sw:P47881 Tr:O70265 Tr:O70266
Tr:O70267 Tr:O70268 Tr:Q62942 Tr:O95918 Tr:O95047"
/codon_start=1
/evidence=not_experimental
/product="dJ1005H11.1 (7 TRANSMEMBRANE RECEPTOR (RHODOPSIN
FAMILY) (OLEFACTORY RECEPTOR LIKE) PROTEIN))"
/protein_id="CAB99212.1"
/db_xref="GI:9368991"
/misc_feature NLHPAKPISFAGRMQTFLESTFAYTECLLVMSYDLVAICHPLRLATMTWRVC
ITLAVTSWTFGVLSLHLVLLPLPFCRPOKITYHFFCEIYALVCIACADTHINMY
LAGAISGLVPLSTIVSYMCILAIQIOSREVORAKRPFCSHLVIGLVGTAII
MYGPRYGNPKRQKYYLLFHSLLFNPMLNPLICSLRNSEVKNLTKRVLGVERAL"
complement(22229..22521)
misc_feature /note="match: STS: Em:H39853"
repeat_region 22385..22489
/note="L2 repeat: matches 2461. .2551 of consensus"
repeat_region 22696..23011
/note="Alusx repeat: matches 1. .312 of consensus"
misc_feature 23044..23513
/note="match: STS: Em:A106560"
complement(23123..23716)
/misc_feature /note="match: STS: Em:AQ276140"
complement(23300..23543)
/misc_feature /note="match: STS: Em:AQ058777"
complement(23751..24305)
/misc_feature /note="match: STS: Em:AQ798220"
complement(23753..24176)
/misc_feature /note="match: STS: Em:AQ550610"
complement(23762..24465)
repeat_region /note="match: STS: Em:B86674"
complement(24220..24343)
/misc_feature /note="FLAM_C repeat: matches 1. .124 of consensus"
complement(24681..25395)
misc_feature /note="match: STS: Em:AQ627093"
complement(24780..25383)
/misc_feature /note="match: STS: Em:AQ627184"
complement(26053..26548)
repeat_region /note="TIGSER1 repeat: matches 1. .521 of consensus"
complement(26549..26861)
repeat_region /note="LIM4 repeat: matches 4025. .4339 of consensus"
complement(26865..27121)
repeat_region /note="Tigger3b repeat: matches 2. .272 of consensus"
complement(27122..27424)
repeat_region /note="Alusx repeat: matches 1. .302 of consensus"
complement(27425..27579)
repeat_region /note="Tigger3b repeat: matches 272. .451 of consensus"
complement(27582..27658)
repeat_region /note="Tigger3b repeat: matches 1155. .1231 of consensus"
complement(27659..29067)
repeat_region /note="LIM4 repeat: matches 2585. .4035 of consensus"
```

```
repeat_region 29189..29287
/note="LIM4 repeat: matches 2499. .2600 of consensus"
misc_feature 29316..29499
/note="match: STS: Em:AA931824"
misc_feature 29507..29990
/note="match: STS: Em:AA931824"
misc_feature 29530..29981
/note="match: STS: Em:AQ356766"
misc_feature 29530..29737
/note="match: STS: Em:AQ253729"
misc_feature 29533..29855
/note="match: STS: Em:AQ319835"
misc_feature 29538..29943
/note="match: STS: Em:AA235111"
misc_feature 29560..29787
/note="match: STS: Em:G27651 Em:T99793"
complement(29779..32112)
/gene="dJ1005H11.2"
complement(29779..32112)
/gene="dJ1005H11.2"
CDS /note="match: proteins: Tr:O95046"
/codon_start=1
/evidence=not_experimental
/product="dJ1005H11.2 (WUGSC_H_DJ0988615.3 PROTEIN)"
/protein_id="CAB99213.1"
/db_xref="GI:9368992"
/translation="MEEPGATPOPYLGLVLELGRVVAALPESMRPDENPYGFPELV"

Query Match 97.0% Score 983.6; DB 9; Length 59130;
Best Local Similarity 99.4% Pred. No. 1.3e-231;
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
```

QY	1	TAAACACTTCTCTAAACCATGAGCATTAATGATTTCTCTGTCATAGGATATGGG	60
Db	22552	TAAACACTTCTCTAAACCATGAGCATTAATGATTTCTCTGTCATAGGATATGGG	22493
QY	61	GACATATATACATCATCAGAGAGTTCCTCTACTGGGATTTCCCGTGGCCCAAGATT	120
Db	22492	GACATATATACATCATCAGAGAGTTCCTCTACTGGGATTTCCCGTGGCCCAAGATT	22433
QY	121	CAGATGCTCCTTTGGGCTCTCTCCCGTCTTACGCTTTCACCTGCGGGAAGGG	180
Db	22432	CAGATGCTCCTTTGGGCTCTCTCCCGTCTTACGCTTTCACCTGCGGGAAGGG	22373
QY	181	ACCATAGTGGGCTGCTTCATGAGTTCAGACTCCAGACTGCAGCGCCCGC-TGTACTTCTCTC	239
Db	22372	ACCATAGTGGGCTGCTTCATGAGTTCAGACTCCAGACTGCAGCGCCCGCATCTTCTCTC	22313
QY	240	TCACACCTGGGGTGTGACATGCGCTTACGCTGCACACAGGTCGCCGATGCTGTG	299
Db	22312	TCACACCTGGGGTGTGACATGCGCTTACGCTGCACACAGGTCGCCGATGCTGTG	22253
QY	300	AACCTCTGATCCAGCCCAAGCCCATCTCTTTGCGGGCGGAGATGACAGACCTTTCTG	359
Db	22252	AACCTCTGATCCAGCCCAAGCCCATCTCTTTGCGGGCGGAGATGACAGACCTTTCTG	22193
QY	360	TTTTCCACTTTTGTGTGACAGAAATGTCTCTCTGCTGTGTGATGCTATGATCTGTAC	419
Db	22192	TTTTCCACTTTTGTGTGACAGAAATGTCTCTCTGCTGTGTGATGCTATGATCTGTAC	22133
QY	420	GTCGCCATCTGACACCCCTCCGATATTTGCCCATCATGACCTGAGAGACTGCATCAC	479
Db	22132	GTCGCCATCTGACACCCCTCCGATATTTGCCCATCATGACCTGAGAGACTGCATCAC	22073
QY	480	CTGCGGTGACTTCTGTCGACACCATGAGAGTCTTTATCTTGATTCATCTGTGTACTT	539
Db	22072	CTGCGGTGACTTCTGTCGACACCATGAGAGTCTTTATCTTGATTCATCTGTGTACTT	22013
QY	540	CTACCTTTACCTCTCTGTAGGCCCCAGAAATTTATCATCTTTTGTGAATCTTGGC	599
Db	22012	CTACCTTTACCTCTCTGTAGGCCCCAGAAATTTATCAC-TTTTGTGGAATCTTGGC	21954
QY	600	TGTTCTCAACTTGCCTGTGACATACCCATCATGAGAACATGGTCTTGCGCGGAGC	659

Db	21953	TGTTCTCAAACTTGCCCTGTGCAGATA	CCACATCAATGAGACATGGTCTTGCCGAGC	21894
QY	660	AATTTCTGGCTGTGGACCCCTGTCCACAT	TGTAGTTTCATATATGTGCATCCTCTG	719
Db	21893	AATTTCTGGGCTGTGGACCCCTGTCCACA	TATGTAGTTTCATATATGTGCATCCTCTG	21834
QY	720	TGCTATCCTTCAGATCCATCAAGGGAAG	TCAGAGGAAGCCCTCTGCACCTGCTTCTC	779
Db	21833	TGCTATCCTTCAGATCCATCAAGGGAAG	TCAGAGGAAGCCCTCTGCACCTGCTTCTC	21774
QY	780	CCACCTCTGTGTGATTTGGACTCTTTTAT	TGGCACAGCCATTCATGTATGTTGGACCCAG	839
Db	21773	CCACCTCTGTGTGATTTGGACTCTTTTAT	TGGCACAGCCATTCATGTATGTTGGACCCAG	21714
QY	840	ATATGGGAACCCCAAGGAGCAGAAGAA	TATCTCCTGCTGTTTCACAGCCTCTTAAATCC	899
Db	21713	ATATGGGAACCCCAAGGAGCAGAAGAA	TATCTCCTGCTGTTTCACAGCCTCTTAAATCC	21654
QY	900	CATGCTCAATCCCCCTTATCTGTAGTCT	TTAGGAACCTCAGAAAGTGAAGATACTTTGAAGAG	959
Db	21653	CATGCTCAATCCCCCTTATCTGTAGTCT	TTAGGAACCTCAGAAAGTGAAGATACTTTGAAGAG	21594
QY	960	AGTCTGGGAGTAGAAAGGCTTTATGMAA	AGATTAAGCATTTGTGACTGACA 1013	
Db	21593	AGTCTGGGAGTAGAAAGGCTTTATGMAA	AGATTAAGCATTTGTGACTGACA 21540	

RESULT	7
AC005587	
LOCUS	AC005587
DEFINITION	Homo sapiens PAC clone RP5-988G15 from 7q33-q35, complete sequence.
ACCESSION	AC005587
VERSION	AC005587.1 GI:4156166
KEYWORDS	HTG.
SOURCE	Homo sapIens.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE
1	(bases 1 to 140915) Hou, S., Woldmann, P. and Le, T.	The sequence of Homo sapiens PAC clone RP5-988G15

REFERENCE	2 (bases 1 to 140513)
AUTHORS	Waterston, R.H.
TITLE	Direct Submission
JOURNAL	Submitted (01-SEP-1998)

REFERENCE	3 (bases 1 to 140915)
AUTHORS	Waterston, R.H.
TITLE	Direct Submission
JOURNAL	Submitted (14-JAN-1999) Genome Sequencing Center, Washington

REFERENCE	4 (bases 1 to 140915)
AUTHORS	Waterston, R.
TITLE	Direct Submission
JOURNAL	Submitted (15-JAN-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE	5 (bases 1 to 140915)

**TITLE** Direct Submission  
**JOURNAL** Submitted (21-DEC-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
**COMMENT** On Jan 14, 1999 this sequence version replaced qi:3907499.

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)  
----- Summary Statistics  
Center project name: H-DJ0988G15

**NOTICE:** This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:**  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

**SOURCE INFORMATION:** This clone was derived from human PAC library RPCI-5, prepared by Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>) using the method described by Ioannou et al., *Nature Genetics* 6:84-9 (1994). The library is from one male donor.

The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong.

NEIGHBORING SEQUENCE INFORMATION:  
Actual start of this clone is at base position 1 of RP5-988G15;  
actual end is at 140915 of RP5-988G15.

FEATURES	Location/Qualifiers
source	1. .140915
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7q33-q35"
	/clone="RP5-988G15"
	/clone.lib="RPC1-5"
	13. .2090
repeat_region	/rpt_family="L1"
	2097. .2259
repeat_region	/rpt_family="L1"
	2300. .2603
repeat_region	/rpt_family="Alu"
	2604. .4120
repeat_region	/rpt_family="L1"
	4123. .4166
repeat_region	/rpt_family="(TA)n"
	4168. .4307
repeat_region	/rpt_family="L1"
	6635. .6668
repeat_region	/rpt_family="AT-rich"
	6829. .7260
repeat_region	/rpt_family="L1"
	7261. .7642
repeat_region	/rpt_family="MaLR"
	7643. .7923
repeat_region	/rpt_family="L1"
	7982. .8283
repeat_region	/rpt_family="Alu"
	8286. .8412
repeat_region	/rpt_family="L1"
	8413. .8708
repeat_region	/rpt_family="Alu"
	8709. .8838
repeat_region	/rpt_family="L1"



```

repeat_region      8840. .8886
/rpt_family="MaLR"
repeat_region      9111. .9139
/rpt_family="AT_rich"
repeat_region      9204. .9490
/rpt_family="Alu"
repeat_region      9646. .9924
/rpt_family="Alu"
repeat_region      10286. .10455
/rpt_family="L2"
repeat_region      10759. .10838
/rpt_family="Mariner"
repeat_region      11127. .11266
/rpt_family="L2"
repeat_region      11850. .12154
/rpt_family="Alu"
repeat_region      12172. .12211
/rpt_family="AT_rich"
repeat_region      12352. .12469
/rpt_family="(TACG)n"
repeat_region      12458. .12607
/rpt_family="(CATA)n"
repeat_region      12580. .12733
/rpt_family="(TA)n"
repeat_region      12660. .12777
/rpt_family="(CATA)n"
repeat_region      12778. .13031
/rpt_family="Alu"
repeat_region      13226. .13535
/rpt_family="L2"
repeat_region      15279. .15520
/rpt_family="MIR"
repeat_region      15954. .16022
/rpt_family="MaLR"
repeat_region      16023. .16332
/rpt_family="MER2_type"
repeat_region      16333. .16371
/rpt_family="MaLR"
repeat_region      16630. .16742
/rpt_family="MaLR"
repeat_region      16816. .17068
/rpt_family="L1"
repeat_region      17088. .17348
/rpt_family="L1"
repeat_region      17857. .17979
/rpt_family="Alu"
repeat_region      18014. .18506
/rpt_family="MaLR"
repeat_region      18707. .19081
/rpt_family="MaLR"
repeat_region      19135. .20717
/rpt_family="Retroviral"
repeat_region      20728. .20748
/rpt_family="AT_rich"
repeat_region      21444. .21599
/rpt_family="MIR"
misc_feature       21688. .22279
/note="60S ribosomal protein L15 pseudogene"
repeat_region      22506. .22540
/rpt_family="AT_rich"
repeat_region      22557. .22583
/rpt_family="AT_rich"
repeat_region      23139. .23162
/rpt_family="(TAA)n"
repeat_region      23696. .24012
/rpt_family="Alu"
repeat_region      24088. .24128
/rpt_family="AT_rich"
repeat_region      24827. .24849
/rpt_family="AT_rich"
repeat_region      24850. .25147
/rpt_family="Alu"
repeat_region      25148. .25173

```

```

repeat_region      25951. .25974
/rpt_family="AT_rich"
repeat_region      26066. .26144
/rpt_family="(CAAA)n"
repeat_region      26388. .26653
/rpt_family="Alu"
repeat_region      27148. .27600
/rpt_family="L1"
repeat_region      28002. .28294
/rpt_family="L2"
repeat_region      /rpt_family="Alu"

Query Match      97.0%; Score 983.6; DB 9; Length 140915;
Best Local Similarity 99.4%; Pred. No. 1.3e-231;
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY      1  TAAACACTTCTCCTAACCATGAGCATTAAGTATGATTTCCCTGTCATAGGATATGGG 60
|||||
Db 101545  TAAACACTTCTCCTAACCATGAGCATTAAGTATGATTTCCCTGTCATAGGATATGGG 101604

QY      61  GACAATATTAACATCCATCAGAGAGTTCCTCTACTGGATTTCCCGTTGGCCCAAGAT 120
|||||
Db 101605  GACAATATTAACATCCATCAGAGAGTTCCTCTACTGGATTTCCCGTTGGCCCAAGAT 101664

QY      121  CAGATGCTCCTCTTTGGGCTCTTCCCTGTTCTACGCTTCACCTGCTGGGAACGG 180
|||||
Db 101665  CAGATGCTCCTCTTTGGGCTCTTCCCTGTTCTACGCTTCACCTGCTGGGAACGG 101724

QY      181  ACCATFACTGGGGCTCATCTCACTGGACATCCAGACTGCACGGCCCC-TGTAATCTTCTC 239
|||||
Db 101725  ACCATFACTGGGGCTCATCTCACTGGACATCCAGACTGCACGGCCCCATGTAATCTTCTC 101784

QY      240  TCACACCTGGCGGTGCTGCACATCGCCCTAGCCCTGCACACAGGTGCCCGGATGCTGTG 299
|||||
Db 101785  TCACACCTGGCGGTGCTGCACATCGCCCTAGCCCTGCACACAGGTGCCCGGATGCTGTG 101844

QY      300  AACCTCTGCATCCAGCCCAAGCCCATCTCTTTGGGGCGGCGATGACAGACCTTTCTG 359
|||||
Db 101845  AACCTCTGCATCCAGCCCAAGCCCATCTCTTTGGGGCGGCGATGACAGACCTTTCTG 101904

QY      360  TTTTCCACTTTTGTGTGCACAGAAATGTCCTCTCTGTTGTTGATGTCCTATGATCTGTAC 419
|||||
Db 101905  TTTTCCACTTTTGTGTGCACAGAAATGTCCTCTCTGTTGTTGATGATCTGTATC 101964

QY      420  GTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCCGAGAGAGTCTGCATCAC 479
|||||
Db 101965  GTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCCGAGAGAGTCTGCATCAC 102024

QY      480  CTCGGGTGACTTCCTGGACCACTGAGTCCTTTATCTCTGATTCATCTGTGTACTT 539
|||||
Db 102025  CTCGGGTGACTTCCTGGACCACTGAGTCCTTTATCTCTGATTCATCTGTGTACTT 102084

QY      540  CTACCTTTAACCTTCTGTAGGCCCCAGAAATTTATCACTTTTGTGAATCTTGGC 599
|||||
Db 102085  CTACCTTTAACCTTCTGTAGGCCCCAGAAATTTATCAC-TTTTGTGAATCTTGGC 102143

QY      600  TGTTCCTCAAACTTGCTGTGAGATACCAATCAATGAGAACATGTCCTGCGCGAGC 659
|||||
Db 102144  TGTTCCTCAAACTTGCTGTGAGATACCAATCAATGAGAACATGTCCTGCGCGAGC 102203

QY      660  AATTCCTGCGGTGTTGGGACCCCTTGTCCACAATTTGATGTTATATATGTGCATCTCTG 719
|||||
Db 102204  AATTCCTGCGGTGTTGGGACCCCTTGTCCACAATTTGATGTTATATATGTGCATCTCTG 102263

QY      720  TGCTATCTTCAGATCCCAATCAAGGAGTTCAGAGGAAGCCCTTGCACCTGCTCTC 779
|||||
Db 102264  TGCTATCTTCAGATCCCAATCAAGGAGTTCAGAGGAAGCCCTTGCACCTGCTCTC 102323

QY      780  CCACCTCTGTGATGAGACTCTTTTATGGCACAGCCATATCATGTATGTGACCCAG 839
|||||
Db 102324  CCACCTCTGTGATGAGACTCTTTTATGGCACAGCCATATCATGTATGTGACCCAG 102383

QY      840  AATGGGAACCCCAAGAGCAGAGAAGAAATATCTCTGCTGTGTTCAACGCTCTTAAATCC 899

```



Db 102384 ATATGGGAACCCCAAGAGCAGAGAATATCTCTGTTTCACAGCCCTTTAATCC 102443  
QY 900 CATGCTCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAGTGAAGATACTTTGAAG 959  
Db 102444 CATGCTCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAGTGAAGATACTTTGAAG 102503  
QY 960 AGTGTGGAGTAGAAGGGCTTATGAAGAAGATTATGCGATTGTGACTGACA 1013  
Db 102504 AGTGTGGAGTAGAAGGGCTTATGAAGAAGATTATGCGATTGTGACTGACA 102557

## RESULT 8

AX078373

LOCUS AX078373 974 bp DNA linear PAT 22-FEB-2001

DEFINITION Sequence 41 from Patent WO0107612.

ACCESSION AX078373

VERSION AX078373.1 GI:13158042

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 974)

Au-Young, J., Bandman, O., Tang, Y.T., Yue, H., Azimzai, Y., Burford, N.,

Baughn, M.R., Lu, D.A., Hillman, J.L., Patterson, C. and Lal, P.

TITLE Receptors and associated proteins

JOURNAL Patent: WO 0107612-A 41 01-FEB-2001;

Incyte Genomics, Inc. (US)

location/Qualifiers

1.974

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/note="Incyte ID No: 2279267CB1"

BASE COUNT 206 a 270 c 214 g 284 t

ORIGIN

Query Match 92.3%; Score 936; DB 6; Length 974;  
Best Local Similarity 99.8%; Pred. No. 5.5e-220;  
Matches 958; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 55 ATGGGGACATATATACATCCATCAGAGTTCCTCTAGGGATTCCCGTTGGCCCA 114  
Db 1 ATGGGGACATATATACATCCATCAGAGTTCCTCTAGGGATTCCCGTTGGCCCA 60  
QY 115 AGGATTGAGATGCTCTTTGGGCTTCTCCCTGTCTTACGCTCTACCCCTGCTGGG 174  
Db 61 AGGATTGAGATGCTCTTTGGGCTTCTCCCTGTCTTACGCTCTACCCCTGCTGGG 120  
QY 175 AACGGACCATCTGGGGCTCATCTGAGTCCAGAGTGCAGAGCCGCC-TCCTACTTC 233  
Db 121 AACGGACCATCTGGGGCTCATCTGAGTCCAGAGTGCAGAGCCGCCCATGCTTTC 180  
QY 234 TTCCTCTCAGACCTGGGGCTGTCGACATCGCCTACGCGCTGCAACAGGTGCCCCGATG 293  
Db 181 TTCCTCTCAGACCTGGGGCTGTCGACATCGCCTACGCGCTGCAACAGGTGCCCCGATG 240  
QY 294 CTGGTGAACCTCTGATCCAGCCAGCCCATCTCTTGGGGCCGCGATGATGACAGC 353  
Db 241 CTGGTGAACCTCTGATCCAGCCAGCCCATCTCTTGGGGCCGCGATGATGACAGC 300  
QY 354 TTTCTGTTTCCACTTTTGTGTGACAGAAATGCTCTCTGCTGTGTGATGCTCTATGAT 413  
Db 301 TTTCTGTTTCCACTTTTGTGTGACAGAAATGCTCTCTGCTGTGTGATGCTCTATGAT 360  
QY 414 CTGTACGTGGCATCTGACACCCCTCCGATATTTGGCCATGACCTGAGAGTCTGC 473  
Db 361 CTGTACGTGGCATCTGACACCCCTCCGATATTTGGCCATGACCTGAGAGTCTGC 420  
QY 474 ATCACCCTGGGGTGAATCTCTGAGCACTGGAGTCTTTATCTCTGATTCATCTTG 533  
Db 421 ATCACCCTGGGGTGAATCTCTGAGCACTGGAGTCTTTATCTCTGATTCATCTTG 480

QY 534 TTACTTCTAACCTTTACCCCTCTGTAGGCCCCAGAGAAATTTATCACTTTTGTGGAAT 593  
Db 481 TTACTTCTAACCTTTACCCCTCTGTAGGCCCCAGAGAAATTTATCACTTTTGTGGAAT 539  
QY 594 CTGGCTGTCTCAAACTTGCCCTGTGCAGATACCCACATCAATGAGACATGCTTGGC 653  
Db 540 CTGGCTGTCTCAAACTTGCCCTGTGCAGATACCCACATCAATGAGACATGCTTGGC 599  
QY 654 CGAGCAATTTCTGGGCTGTGGGACCCCTTGTCCACAATTTAGTTTATATATGTGCAT 713  
Db 600 CGAGCAATTTCTGGGCTGTGGGACCCCTTGTCCACAATTTAGTTTATATATGTGCAT 659  
QY 714 CCTCTGTCTATCCTTCAGATCCAAATCAAGGGAAGTTGAGGAAAGCCCTTGTGCACCTG 773  
Db 660 CCTCTGTCTATCCTTCAGATCCAAATCAAGGGAAGTTGAGGAAAGCCCTTGTGCACCTG 719  
QY 774 CTCTGCCACCTCTGTGTGATGAGCTCTTTATGCGACACCCATTTATCATGTATGTTGG 833  
Db 720 CTCTGCCACCTCTGTGTGATGAGCTCTTTATGCGACACCCATTTATCATGTATGTTGG 779  
QY 834 ACCGAGATATGGGAACCCCAAGGACGACAGAGAAATATCTCTGCTTTACAGCCCTCTT 893  
Db 780 ACCGAGATATGGGAACCCCAAGGACGACAGAGAAATATCTCTGCTTTACAGCCCTCTT 839  
QY 894 TAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGGAACTCAGAGTGAAGATACTTT 953  
Db 840 TAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGGAACTCAGAGTGAAGATACTTT 899  
QY 954 GAAGAGAGTCTGGGAGTAGAAGGGCTTTATGAAGAAGATTATGGCATTTGACTGACA 1013  
Db 900 GAAGAGAGTCTGGGAGTAGAAGGGCTTTATGAAGAAGATTATGGCATTTGACTGACA 959

## RESULT 9

AX242191

LOCUS AX242191

DEFINITION Sequence 939 from Patent WO0127158.

ACCESSION AX242191

VERSION AX242191.1 GI:15799066

KEYWORDS

SOURCE

ORGANISM

synthetic construct.

artificial sequences.

REFERENCE 1 (bases 1 to 930)

AUTHORS Bellenson, J., Smith, D., Lancet, D., Glusman, G., Fuchs, T. and

Yanai, I.

TITLE Olfactory receptor sequences

JOURNAL Patent: WO 0127158-A 939 19-APR-2001;

Dislicents (US); YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)

location/Qualifiers

1.930

/organism="synthetic construct"

/db\_xref="taxon:32630"

/note="(H389789 nucleotide)"

BASE COUNT 192 a 266 c 201 g 271 t

ORIGIN

Query Match 89.3%; Score 905.4; DB 6; Length 930;  
Best Local Similarity 99.7%; Pred. No. 1.9e-212;  
Matches 928; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 55 ATGGGGACATATATACATCCATCAGAGTTCCTCTAGGGATTCCCGTTGGCCCA 114  
Db 1 ATGGGGACATATATACATCCATCAGAGTTCCTCTAGGGATTCCCGTTGGCCCA 60  
QY 115 AGGATTGAGATGCTCTTTGGGCTTCTCCCTGTCTTACGCTCTACCCCTGCTGGG 174  
Db 61 AGGATTGAGATGCTCTTTGGGCTTCTCCCTGTCTTACGCTCTACCCCTGCTGGG 120  
QY 175 AACGGACCATCTGGGGCTCATCTGAGTCCAGAGTGCAGAGCCGCC-TCCTACTTC 233  
Db 121 AACGGACCATCTGGGGCTCATCTGAGTCCAGAGTGCAGAGCCGCCCATGCTTTC 180

QY	234	TTCCCTCACACCTGGCGGTGCTGCACATCGCCTACGGCTGCACACACGGTGGCCCGGATG	293
	181	TTCCCTCACACCTGGCGGTGCTGCACATCGCCTACGGCTGCACACACGGTGGCCCGGATG	240
QY	294	CTGGTGAACCTCTCTGCATTCACAGCCCAAGCCCATCTCTTTGGCGGGCCGATGATGCAGACC	353
Db	241	CTGGTGAACCTCTCTGCATTCACAGCCCAAGCCCATCTCTTTGGCGGGCCGATGATGCAGACC	300
QY	354	TTTCTGTTTTCACACTTTTGCTGTCAACAGAAATGTCCTCTCTGCTGGTGTATGTCCTATGAT	413
Db	301	TTTCTGTTTTCACACTTTTGCTGTCAACAGAAATGTCCTCTCTGCTGGTGTATGTCCTATGAT	360
QY	414	CTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC	473
Db	361	CTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC	420
QY	474	ATCACCCCTCGCGGTGACTTCTCTGACACCACTGAGTCCTTTTATCCTTGATTCATCTGTG	533
Db	421	ATCACCCCTCGCGGTGACTTCTCTGACACCACTGAGTCCTTTTATCCTTGATTCATCTGTG	480
QY	534	TTACTTCTACCTTTACCCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTGTGTAAT	593
Db	481	TTACTTCTACCTTTACCCCTTCTGTAGGCCCCAGAAAATTTATCACTTTTGTGTAAT	539
QY	594	CTTGGCTGTCTCAAACTTGCTGTGACATACCCACATCAATGAGAACATGGTCTTGCC	653
Db	540	CTTGGCTGTCTCAAACTTGCTGTGACATACCCACATCAATGAGAACATGGTCTTGCC	599
QY	654	CGGAGCAATTTCTGGGCTGGTGGGACCCCTGTCCACAAATGTAGTTTCAATATATGTGCAT	713
Db	600	CGGAGCAATTTCTGGGCTGGTGGGACCCCTGTCCACAAATGTAGTTTCAATATATGTGCAT	659
QY	714	CCTCTGTGCTATCCTTCAGATCCAAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG	773
Db	660	CCTCTGTGCTATCCTTCAGATCCAAATCAAGGGAAGTTCAGAGGAAAGCCTTCTGCACCTG	719
QY	774	CTTCTCCACCTCTGTGTGATTTGGACTCTTTATGGCACAGCCATTATCATGTATGTTGG	833
Db	720	CTTCTCCACCTCTGTGTGATTTGGACTCTTTATGGCACAGCCATTATCATGTATGTTGG	779
QY	834	ACCCAGATATGGGAACCCCAAGAGCAGAGAATATCTCCTGCTGTTTACAGACCTCTT	893
Db	780	ACCCAGATATGGGAACCCCAAGAGCAGAGAATATCTCCTGCTGTTTACAGACCTCTT	839
QY	894	TATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAAGCTCAGAAAGTGAAGATACTTT	953
Db	840	TATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAAGCTCAGAAAGTGAAGATACTTT	899
QY	954	GAAGAGAGTGTGGAGTAGAAGGCGCTTTA	984
Db	900	GAAGAGAGTGTGGAGTAGAAGGCGCTTTA	930
RESULT	10		
LOCUS	AX451262	933 bp	DNA linear PAT 03-JUL-2002
DEFINITION	Sequence 1455 from Patent WO0214501.		
ACCESSION	AX451262		
VERSION	AX451262.1	GI:21698331	
KEYWORDS	unidentified.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1		
AUTHORS	Hubert,R.S., Raitano,A.B., Faris,M., Challita-Eid,P.M., Ge,W. and Jakobovits,A.		
TITLE	Nucleic acids and corresponding proteins entitled phori-ai1 and phori-f5d6 useful in treatment and detection of cancer		
JOURNAL	Patent: WO 0214501-A 1455 21-FEB-2002;		
FEATURES	Agensys, Inc. (US)		
source	1.933	Location/Qualifiers	
	/organism="unidentified"		

CDS		/db_xref="taxon:32644"	/note="PHORI-F5D6"	1..933	/note="unnamed protein product"	/codon_start=1	/protein_id="CAD37777.1"	/db_xref="GI:21698332"	LGLSIDSRLLHAPMEFELSHLAVVDIAYACNTVRMLVNLHPAKIPISFAGRMQTEFF ESTFAVTECLLVMSYDLVAICHLPRYLAINTRVCITLATVSWTGTLSLIHLY LLPLPECRPKRIYHECEILAVLKLCADRYTHINENVLGASIGLVGPLSTIVSYN CIIICALIQISREYORKAFRTCSFLCVIGLVGFALIMVGPRGNPKQKKYILLE HSLENPMLNPLICSIRNSEVKNTIKRVLGERAL"
BASE COUNT	194	a	265	c	203	g	271	t	
ORIGIN									
Query Match		89.1%	Score 903.6;	DB 6;	Length 933;				
Best Local Similarity		99.4%;	Pred. No.5.4e-212;						
Matches 928;	Conservative	0;	Mismatches 4;	Indels 2;	Gaps 2;				
OY	55	ATGGGSGACAATATATACATCCATCACAGAGTTCTCTACTGGGATTCCCGTTGGCCA	114						
Db	1	ATGGAGACAATATATAACATCCATCAGAGACTTCCTCTACTGGGATTCCCGTTGGCCA	60						
OY	115	AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTCTACGTTCTACCCCTGCTGGGG	174						
Db	61	AGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTCTACGTTCTACCCCTGCTGGGG	120						
OY	175	AACGGGACCATACTGGGGCTCATCTCACTGGAGCTCCAGACTGCACGCCGCC--TGTA	233						
Db	121	AACGGGACCATACTGGGGCTCATCTCACTGGAGCTCCAGACTGCACGCCGCCATGTA	180						
OY	234	TTCCTTCACACCTGGCGGTGCTGCACATCGCCCTACGCCCTGAACACGGTGGCCCCG	293						
Db	181	TTCCTTCACACCTGGCGGTGCTGCACATCGCCCTACGCCCTGAACACGGTGGCCCCG	240						
OY	294	CTGGTAGA	353						
Db	241	CTGGTAGA	300						
OY	354	TTTTCTGTTTCCACTTTTGTCTGTACAGAAATGTCCTCTCTGTGTGTATGTCCTAT	413						
Db	301	TTTTCTGTTTCCACTTTTGTCTGTACAGAAATGTCCTCTCTGTGTGTATGTCCTAT	360						
OY	414	CTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC	473						
Db	361	CTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC	420						
OY	474	ATCACCCCTCGGGTGACTTCTCTGGACCACTGAGTCTTTATCCTTGATTCATCTTGTG	533						
Db	421	ATCACCCCTCGGGTGACTTCTCTGGACCACTGAGTCTTTATCCTTGATTCATCTTGTG	480						
OY	534	TYACCTTACCTTTTACCCCTCTGTAGGCCCCAGAAAATTTATACATTTT	593						
Db	481	TYACCTTACCTTTTACCCCTCTGTAGGCCCCAGAAAATTTATAC - TTTT	539						
OY	594	CTTGGCTGTCTCAAACCTTGCCCTGTGCAGATACCCACATCAATGAGA	653						
Db	540	CTTGGCTGTCTCAAACCTTGCCCTGTGCAGATACCCACATCAATGAGA	599						
OY	654	CGGAGCAATTTCTGGGCTGGTGGGACCCCTGTCCACAATTTAGTTTCATATATGTGCAT	713						
Db	600	CGGAGCAATTTCTGGGCTGGTGGGACCCCTGTCCACAATTTAGTTTCATATATGTGCAT	659						
OY	714	CCTCTGTGCTATCTTCAGATCCATCAAGGGAAGTTTCAGAGGAAAGCCTTCTGCACCTG	773						
Db	660	CCTCTGTGCTATCTTCAGATCCATCAAGGGAAGTTTCAGAGGAAAGCCTTCTGCACCTG	719						
OY	774	CTTCTCCACCTCTGTGTGATTTGAGCTCTTTATGGACAGCCATTTATCATGTATGTGG	833						
Db	720	CTTCTCCACCTCTGTGTGATTTGAGCTCTTTATGGACAGCCATTTATCATGTATGTGG	779						
OY	834	ACCCAGATATGGGAACCCCAAGGAGCAGAGAAGAAATATCTCTGCTTTCACAGCCTCTT	893						

```
|||||
Db 780 ACCGATATGGGAACCCCAAGAGACAGAGAATAATCTCTGCTGTTTCACAGCCTCTT 839
QY 894 TAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAACTCAGAAAGTGAAGATACTTT 953
|||||
Db 840 TAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAACTCAGAAAGTGAAGATACTTT 899
QY 954 GAAGAGAGTGTGGAGTAGAAGGGCTTTATGA 987
|||||
Db 900 GAAGAGAGTGTGGAGTAGAAGGGCTTTATGA 933

RESULT 11
AX241442 930 bp DNA linear PAT 26-SEP-2001
LOCUS AX241442
DEFINITION Sequence 190 from Patent WO0127158.
ACCESSION AX241442
VERSION AX241442.1 GI:15798317
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 930)
AUTHORS Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
Yanal,I.
TITLE Olfactory receptor sequences
JOURNAL Patent: WO 0127158-A 190 19-APR-2001;
Discent (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES
source 1..930
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="(H38938 nucleotide)"

BASE COUNT 193 a 265 c 202 g 270 t
ORIGIN

Query Match 88.8%; Score 900.6; DB 6; Length 930;
Best Local Similarity 99.4%; Pred. No. 2.9e-211;
Matches 925; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 55 ATGGGGACAATATAACATCCATCACAAGATTCCTCTACTGGGATTCCTGGCCCA 114
|||||
Db 1 ATGGGAGACAATATAACATCCATCACAAGATTCCTCTACTGGGATTCCTGGCCCA 60
QY 115 AGGATTCAGATGCTCTCTTGGGCTCTTCTCCCTGTCTTACGTTCTCACCCGCTGGGG 174
|||||
Db 61 AGGATTCAGATGCTCTCTTGGGCTCTTCTCCCTGTCTTACGTTCTCACCCGCTGGGG 120
QY 175 AACGGACCATACTGGGGCTCACTCACTGAGCTCAGACTGCAGCCGCC-TGTACTTC 233
|||||
Db 121 AACGGACCATACTGGGGCTCACTCACTGAGCTCAGACTGCAGCCGCCATGTACTTC 180
QY 234 TTCCTTCACACCTGGGGCTCGTGCATCGCTACGCTGCAACAGGTCGCCCGGATG 293
|||||
Db 181 TTCCTTCACACCTGGGGCTCGTGCATCGCTACGCTGCAACAGGTCGCCCGGATG 240
QY 294 CTGTGACCTCTGCATCGACCAAGCCCATCTCTTGGGGGCCGCGCATGATGACAGACC 353
|||||
Db 241 CTGTGACCTCTGCATCGACCAAGCCCATCTCTTGGGGGCCGCGCATGATGACAGACC 300
QY 354 TTTCTGTTTCCACTTTTGTCTGTACAGATGTCTCTCCCTGTGTGTGATGTCTATGAT 413
|||||
Db 301 TTTCTGTTTCCACTTTTGTCTGTGTACAGATGTCTCTCCCTGTGTGTGATGTCTATGAT 360
QY 414 CTGTAGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGC 473
|||||
Db 361 CTGTAGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGC 420
QY 474 ATCACCCCTGGGGTGAATCTCTGGACCACTGAGTCTTTATCTCTGATTCATCTTGTG 533
|||||
Db 421 ATCACCCCTGGGGTGAATCTCTGGACCACTGAGTCTTTATCTCTGATTCATCTTGTG 480
QY 534 TTACTTCTACCTTTACCCCTCTGTAGGCCCCAGAAATTATCACTTTTGTGTGAAT 593
```

```
|||||
Db 481 TTACTTCTACCTTTACCTTCTGTAGGCCCCAGAAATTATAC-TTTTTTGTGAAT 539
QY 594 CTGGCTGTCTCAAACTTGCTGTGAGATACCCATCAATGAGACATGCTTGGC 653
|||||
Db 540 CTGGCTGTCTCAAACTTGCTGTGAGATACCCATCAATGAGACATGCTTGGC 599
QY 654 CGAGCAATTTCTGGGCTGGTGAGCCCTTGTCCACAATTTAGTTTATATATGTGCAT 713
|||||
Db 600 CGAGCAATTTCTGGGCTGGTGAGCCCTTGTCCACAATTTAGTTTATATATGTGCAT 659
QY 714 CTTCTGTGCTATCTTCAGATCCATCAATCAAGGAGTTTCAGAGGAAACCTTCTGCACCTG 773
|||||
Db 660 CTTCTGTGCTATCTTCAGATCCATCAATCAAGGAGTTTCAGAGGAAACCTTCTGCACCTG 719
QY 774 CTTCTCCACCTCTGTGTGATTTGACTCTTTATGACACAGCCATTCATGTATGTGG 833
|||||
Db 720 CTTCTCCACCTCTGTGTGATTTGACTCTTTATGACACAGCCATTCATGTATGTGG 779
QY 834 ACCGATATGGGAACCCCAAGAGACAGAGAATAATCTCTGCTGTTTCACAGCCTCTT 893
|||||
Db 780 ACCGATATGGGAACCCCAAGAGACAGAGAATAATCTCTGCTGTTTCACAGCCTCTT 839
QY 894 TAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAACTCAGAAAGTGAAGATACTTT 953
|||||
Db 840 TAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAACTCAGAAAGTGAAGATACTTT 899
QY 954 GAAGAGAGTGTGGAGTAGAAGGGCTTTA 984
|||||
Db 900 GAAGAGAGTGTGGAGTAGAAGGGCTTTA 930

RESULT 12
AX241473 930 bp DNA linear PAT 26-SEP-2001
LOCUS AX241473
DEFINITION Sequence 221 from Patent WO0127158.
ACCESSION AX241473
VERSION AX241473.1 GI:15798348
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 930)
AUTHORS Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and
Yanal,I.
TITLE Olfactory receptor sequences
JOURNAL Patent: WO 0127158-A 221 19-APR-2001;
Discent (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES
source 1..930
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="(H38970 nucleotide)"

BASE COUNT 193 a 265 c 202 g 270 t
ORIGIN

Query Match 88.8%; Score 900.6; DB 6; Length 930;
Best Local Similarity 99.4%; Pred. No. 2.9e-211;
Matches 925; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 55 ATGGGGACAATATAACATCCATCACAAGATTCCTCTACTGGGATTTCCCGTTGGCCCA 114
|||||
Db 1 ATGGGAGACAATATAACATCCATCACAAGATTCCTCTACTGGGATTTCCCGTTGGCCCA 60
QY 115 AGGATTCAGATGCTCTCTTGGGCTCTTCTCCCTGTCTTACGTTCTCACCCGCTGGGG 174
|||||
Db 61 AGGATTCAGATGCTCTCTTGGGCTCTTCTCCCTGTCTTACGTTCTCACCCGCTGGGG 120
QY 175 AACGGACCATACTGGGGCTCACTCACTGAGCTCAGACTGCAGCCGCC-TGTACTTC 233
|||||
Db 121 AACGGACCATACTGGGGCTCACTCACTGAGCTCAGACTGCAGCCGCCATGTACTTC 180
QY 234 TTCCTTCACACCTGGGGCTCGTGCATCGCTACGCTGCAACAGGTCGCCCGGATG 293
```

|||||  
Db 181 TTCTCTACACCTGGCGGTGCTGCACATCGCCCTGACACAGGTGCCCCGATG 240  
QY 294 CTGTGACACCTCCATCCATCCAGCCAGCCATCTCTTGGGGCCGATGACAGACC 353  
Db 241 CTGTGACACCTCCATCCATCCAGCCAGCCATCTCTTGGGGCCGATGACAGACC 300  
QY 354 TTTCTGTTTCCACTTTTGTCTGTACAGAAATGCTCTCTGCTGTGTATGCTATGAT 413  
Db 301 TTTCTGTTTCCACTTTTGTCTGTACAGAAATGCTCTCTGCTGTGTATGCTATGAT 360  
QY 414 CTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 473  
Db 361 CTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGGAGAGTCTGC 420  
QY 474 ATCACCCTGGGGTGAATCTCTGGACCACTGGAGTCTTTATCTCTGATTCATCTTGTG 533  
Db 421 ATCACCCTGGGGTGAATCTCTGGACCACTGGAGTCTTTATCTCTGATTCATCTTGTG 480  
QY 534 TTACTTCTACCTTTACCCCTTCTGTAGGCCCCAGAAATTTATCCTTTTGTGAAT 593  
Db 481 TTACTTCTACCTTTACCCCTTCTGTAGGCCCCAGAAATTTATCCTTTTGTGAAT 539  
QY 594 CTGGCTGTCTCAAACTTGCCCTGTGCAGATACCCATCATGAGAACATGGTCTTGGC 653  
Db 540 CTGGCTGTCTCAAACTTGCCCTGTGCAGATACCCATCATGAGAACATGGTCTTGGC 599  
QY 654 CGGAGCAATTTCTGGGCTGGTGGGACCCCTTGCCACAATTTAGTTTCAATATATGTGCAT 713  
Db 600 CGGAGCAATTTCTGGGCTGGTGGGACCCCTTGCCACAATTTAGTTTCAATATATGTGCAT 659  
QY 714 CCTCTGTCTATCTTCATGATCCATCAAGGGAAGTTTCAGAGAAAGCCTTCTGCACCTG 773  
Db 660 CCTCTGTCTATCTTCATGATCCATCAAGGGAAGTTTCAGAGAAAGCCTTCTGCACCTG 719  
QY 774 CTCTCCACCTCTGTGTATGATGAGACTCTTTATGACACAGCCATTCATGATGTTGG 833  
Db 720 CTCTCCACCTCTGTGTATGATGAGACTCTTTATGACACAGCCATTCATGATGTTGG 779  
QY 834 ACCCAGATATGGGAACCCCAAGAGCAGAGAAATATCTCTGCTGTTTCACAGCCCTCT 893  
Db 780 ACCCAGATATGGGAACCCCAAGAGCAGAGAAATATCTCTGCTGTTTCACAGCCCTCT 839  
QY 894 TAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGATACTTT 953  
Db 840 TAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGGAACTCAGAAGTGAAGATACTTT 899  
QY 954 GAAGAGAGTGTGGAGTAGAAGGGCTTTA 984  
Db 900 GAAGAGAGTGTGGAGTAGAAGGGCTTTA 930  
RESULT 13  
LOCUS AC091768 104955 bp DNA linear PRI 09-JAN-2002  
DEFINITION Homo sapiens BAC clone RP11-703N5 from 7, complete sequence.  
ACCESSION AC091768  
VERSION AC091768.4 GI:15145624  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Sulston, J.E. and Waterston, R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
MEDLINE 99063792  
PUBMED 9847074  
REFERENCE 2 (bases 1 to 104955)  
AUTHORS Harris, A., Haakenson, W. and Spalding, L.  
TITLE The sequence of Homo sapiens BAC clone RP11-703N5  
JOURNAL Unpublished (2001)

REFERENCE 3 (bases 1 to 104955)  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (31-MAY-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 4 (bases 1 to 104955)  
AUTHORS Waterston, R.H.  
TITLE Direct Submission  
JOURNAL Submitted (09-AUG-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 5 (bases 1 to 104955)  
AUTHORS Waterston, R.  
TITLE Direct Submission  
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Aug 9, 2001 this sequence version replaced gi:14550325.  
COMMENT  
-----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc  
Contact: sapiens@wustl.wustl.edu  
----- Summary Statistics  
Center project name: H\_NH0703N05  
-----  
NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.  
  
This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.  
  
MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see  
http://www.nhgri.nih.gov/DIR/CTB/CHR7, send  
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc  
  
SOURCE INFORMATION:  
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (http://bacpac.med.buffalo.edu)  
VECTOR: pBac3.6  
NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP4-669B10, 2000 bp overlap; the clone sequenced to the right is RP4-798C17, 2000 bp overlap.  
Actual start of this clone is at base position 123947 of RP4-669B10; actual end is at base position 64342 of RP11-798C17.  
RP11-703N5 contains a single plasmid region from 4530 to 4616 and 4668 to 4721.  
FEATURES  
source location/qualifiers  
1..104955  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="7"



```
/map="7"
/clone="RP11-703N5"
/clone_lib="RPCT-11"
1..162
/rpt_family="L1"
repeat_region
/rpt_family="AT_rich"
928..1614
/rpt_family="L2"
repeat_region
2091..2204
/rpt_family="L2"
repeat_region
2649..2713
/rpt_family="ERV1"
repeat_region
2651..2710
/rpt_family="(CA)n"
repeat_region
4071..4108
/rpt_family="(TCCA)n"
repeat_region
4472..4511
/rpt_family="L2"
misc_feature
5679..5892
/note="similar to Homo sapiens EST AA298029
(NID:g1950363)"
7802..7825
/rpt_family="AT_rich"
repeat_region
8020..8173
/rpt_family="MIR"
repeat_region
9639..9865
/rpt_family="CR1"
repeat_region
10327..10490
/rpt_family="Alu"
repeat_region
10476..10498
/rpt_family="(CAA)n"
repeat_region
11042..11088
/rpt_family="GA-rich"
repeat_region
11387..11406
/rpt_family="(T)n"
repeat_region
11947..12268
/rpt_family="L1"
misc_feature
12489..13017
/note="match to EST A1478847 (NID:g4372015) tm24e12.x1"
12793..12973
/rpt_family="L1"
repeat_region
12974..13192
/rpt_family="MalR"
misc_feature
13147..13148
/note="match to EST A1478847 (NID:g4372015) tm24e12.x1"
13193..13485
/rpt_family="Alu"
repeat_region
13486..13668
/rpt_family="MalR"
repeat_region
13669..14177
/rpt_family="MalR"
repeat_region
14177..15248
/rpt_family="MalR"
repeat_region
15249..16333
/rpt_family="ERVk"
repeat_region
16334..16622
/rpt_family="MalR"
repeat_region
16628..16988
/rpt_family="MalR"
repeat_region
16989..17116
/rpt_family="MalR"
repeat_region
17119..17273
/rpt_family="L1"
repeat_region
18350..18588
/rpt_family="ERVl"
repeat_region
18658..18733
/rpt_family="ERVl"
repeat_region
19319..19982
/rpt_family="L1"
repeat_region
20126..20161
/rpt_family="AT_rich"
repeat_region
20389..21008
```

```
/rpt_family="L1"
repeat_region
21009..23092
/rpt_family="L1"
repeat_region
22685..22728
/rpt_family="(TG)n"
repeat_region
23114..23443
/rpt_family="L1"
repeat_region
23523..23592
/rpt_family="AT_rich"
repeat_region
25402..26143
/rpt_family="L1"
repeat_region
26433..26513
/rpt_family="CR1"
repeat_region
27066..27227
/rpt_family="MER1_type"
repeat_region
27289..27962
/rpt_family="L1"
repeat_region
27954..28022
/rpt_family="AT_rich"
misc_feature
28774..29234
/note="similar to Homo sapiens EST BF116115
(NID:g10985591)"
28774..29205
/note="similar to Homo sapiens EST BE179276
(NID:g8658428)"
28802..29204
/note="similar to Homo sapiens EST H62441 (NID:g1015273)
y79d08.r1"
28816..29238
/note="similar to Homo sapiens EST AI079550 (NID:g3415801)
O204f08.x1"
28816..29238
/note="similar to Homo sapiens EST A1148854 (NID:g3677323)
qc70a02.x1"
28816..29238

Query Match 66.6%; Score 675; DB 9; Length 104955;
Best Local Similarity 83.0%; Pred.No.1.6e-155;
Matches 793; Conservative 0; Mismatches 160; Indels 2; Gaps 2;

QY 36 TTTCCTGTGATAGGATATGCGGACATATATACATCATCAGAGTTCCTACT 95
Db 28775 TTGTTCTTCTACAGGAAATGCGGAAATCAGACTTCATCAGAGTTCCTACT 28834

QY 96 GGGATTTCCCGTTGGCCCAAGATTCAGATGCTCTTGGGCTCTTCCCTGTTCTA 155
Db 28835 GGGATTTCCCGATGGCCCAAGATTCAGATGCTCTTGGGCTCTTCCCTGTTCTA 28894

QY 156 CGTCTTCAACCTGTGGGGAACGGACATCTGGGGCTCATCTCAGTGCAGACT 215
Db 28895 CATCTTCAATCTGTGGGAACGGGACATCTGGGGCTCATCTCAGTGCAGACT 28954

QY 216 GCACGCCCC-7GTACTTCTTCTCCTCAGACCTGGCGGTGACATGCGCTTGC 274
Db 28955 CCACACCCCATGACTTCTTCTCCTCAGACCTGGCGGTGACATGCGCTTGC 29014

QY 275 CAACACGGTCCCGGATGCTGTGAACCTCTGCATCCAGCCAGCCATCTCTTGC 334
Db 29015 CAGCAGGGTGCCCAAGATGCTGTGAACCTCTGCATCCAGCCAGCCATCTCTTGC 29074

QY 335 GGGCCGATGATGACAGACCTTCTGTTCACCTTTTGTGCTGACAGATGCTCTCT 394
Db 29075 TGCGTCATGACCCAGATGTTTCTGTGTTTGTGAGTTTGTGACATGATGCTCTCT 29134

QY 395 GGTGTGATGTCATATGATCTGTACGTGCGCATCTGCCACCCCTCGATATTTGGCAT 454
Db 29135 GGTGTGATGTCATATGATCGGTACGTGCGCATCTGCCACCCCTCGATATTTACCAT 29194

QY 455 CATGACCTGGAGAGTCTGCATCACCCTGCGGTGACTTCTGTGACCACTGGAGTCTTT 514
Db 29195 CATGACCTGGAGAGTCTGCATCACCCTGCGGTGACTTCTGTGACCACTGGAGTCTTT 29254

QY 515 ATCCTGATTCATCTGTGTACTTCTTACCTTTACCCCTCTGTAGGCCCCAGAAATTTA 574
```



Db 29255 GGGCCCTGTCCATCTAGTGTACTGCTACACACTGCTCTGTGGACCCAGAACTTAA 29314  
OY 575 TCACATTTTGTGAAATCTGGCTGTCTCAACTTGGCTGTGCAGATACCCACATCA 634  
Db 29315 TCAC-TTTTCTGTGAAATATGACTGTCTCAACTTGGCTGTGCAGATACCCACATTA 29373  
OY 635 ATGAGACATGGTCTTGGCCGAGCAATTTCTGGGCTGGTGAGCCCTTGTCCACAAATTG 694  
Db 29374 ATGAGGTAAATGGTTTGGCAGGGGCAAGTGTCTGTGGTGAGCCCTTGTCTTCCACTG 29433  
OY 695 TAGTTTCATATATGTGCATCCTCTGTGCTATCCTTCAGATCCAAATCAAGGAGTTCAGA 754  
Db 29434 TAATATCTTATGTTCATATCTATGTGCCATTTCTAAAGATCCAGTCAGAGAGGGGTGCC 29493  
OY 755 GGAAGCCTTCTGCACCTGCTTCTCCACCTCTGTGTGATTTGACTCTTTTATGGCACAG 814  
Db 29494 AGAAGCCTTCTCCATCTGCTCTCCACCTCTGTGTGATTTGACTCTTTATGGCACAG 29553  
OY 815 CCATATCATGTATGTGTGACCCAGATATGGAACCCCAAGAGAGAGAAATATCTCC 874  
Db 29554 CCATCATCATGTATGTGTGAGCCCAAGATATGAGAGCCCAAGAGAGAGAAATATCTCC 29613  
OY 875 TGCCTTTCACAGCCTCTTAAATCCCATGCTCAATCCCTTATCTGTAGCTTAGAACT 934  
Db 29614 TGCCTTTCACAGCCTCTTCAATCCCATGCTTAAATCCCTTATCTGTAGCTTAGAACT 29673  
OY 935 CAGAAGTGAAGATACTTTGAAGAGAGAGTGTGGGAGTAGAAAGGGCTTTATGAAA 989  
Db 29674 AGGAAGTCCAAGGTACTCTAAAGAGAGATGCTTGAAGAAGAGAACTTCATGAAA 29728

RESULT 14  
AC091746 204316 bp DNA linear HTG 26-APR-2002  
LOCUS AC091746  
DEFINITION Mus musculus chromosome 6 clone RP23-62E15 map 6B2, \*\*\* SEQUENCING  
IN PROGRESS \*\*\*, 30 ordered pieces.  
ACCESSION AC091746 GI:20331040  
VERSION AC091746.2  
KEYWORDS HTG; HTGS\_PHASE2.  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 204316)  
AUTHORS Young,J.M., Madan,A., Hall,J., Friedman,C., Dickhoff,R., Greene,J.,  
Lane,R.P. and Trask,B.J.  
TITLE Evolution of the mouse olfactory receptor gene family  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 204316)  
AUTHORS Young,J.M., Madan,A., Hall,J., Friedman,C., Dickhoff,R., Greene,J.,  
Lane,R.P. and Trask,B.J.  
TITLE Direct Submission  
JOURNAL Submitted (28-MAY-2001) Division of Human Biology, Fred Hutchinson  
Cancer Research Center, 1100 Fairview Ave N, C3-168, P.O. Box  
19024, Seattle, WA 98109-1024, USA  
REFERENCE 3 (bases 1 to 204316)  
AUTHORS Young,J.M., Madan,A., Hall,J., Friedman,C., Dickhoff,R., Greene,J.,  
Rodrigues,S., Ketterman,M., Faney,J., Helton,E., Sanchez,A.,  
Madan,A., Whiting,M., Ross,J.A., Lane,R.P. and Trask,B.J.  
TITLE Direct Submission  
JOURNAL Submitted (26-APR-2002) Division of Human Biology, Fred Hutchinson  
Cancer Research Center, 1100 Fairview Ave N, C3-168, P.O. Box  
19024, Seattle, WA 98109-1024, USA  
COMMENT On Apr 26, 2002 this sequence version replaced gi:14211553.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 30 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.  
1 11127: contig of 11127 bp in length  
\* 11128 11227: gap of unknown length  
\* 11228 18868: contig of 7641 bp in length  
\* 18869 18968: gap of unknown length  
\* 18969 40960: contig of 21992 bp in length  
\* 40961 41060: gap of unknown length  
\* 41061 44591: contig of 3531 bp in length  
\* 44592 44691: gap of unknown length  
\* 44692 46067: contig of 1376 bp in length  
\* 46068 46167: gap of unknown length  
\* 46168 47511: contig of 1344 bp in length  
\* 47512 47611: gap of unknown length  
\* 47612 51354: contig of 3743 bp in length  
\* 51355 51454: gap of unknown length  
\* 51455 52639: contig of 1185 bp in length  
\* 52640 52739: gap of unknown length  
\* 52740 54199: contig of 1460 bp in length  
\* 54200 54299: gap of unknown length  
\* 54300 63193: contig of 8894 bp in length  
\* 63194 63293: gap of unknown length  
\* 63294 70625: contig of 7332 bp in length  
\* 70626 70725: gap of unknown length  
\* 70726 74065: contig of 3340 bp in length  
\* 74066 74165: gap of unknown length  
\* 74166 77890: contig of 3725 bp in length  
\* 77891 77990: gap of unknown length  
\* 77991 82776: contig of 4786 bp in length  
\* 82777 82876: gap of unknown length  
\* 82877 90955: contig of 8079 bp in length  
\* 90956 91055: gap of unknown length  
\* 91056 93563: contig of 2508 bp in length  
\* 93564 93663: gap of unknown length  
\* 93664 98539: contig of 4876 bp in length  
\* 98540 98639: gap of unknown length  
\* 98640 100233: contig of 1594 bp in length  
\* 100234 100333: gap of unknown length  
\* 100334 108744: contig of 8411 bp in length  
\* 108745 108844: gap of unknown length  
\* 108845 120584: contig of 11740 bp in length  
\* 120585 120684: gap of unknown length  
\* 120685 122234: contig of 1550 bp in length  
\* 122235 122334: gap of unknown length  
\* 122335 136966: contig of 14632 bp in length  
\* 136967 137066: gap of unknown length  
\* 137067 146815: contig of 9749 bp in length  
\* 146816 146915: gap of unknown length  
\* 146916 152543: contig of 5628 bp in length  
\* 152544 152643: gap of unknown length  
\* 152644 164250: contig of 11607 bp in length  
\* 164251 164350: gap of unknown length  
\* 164351 170223: contig of 5873 bp in length  
\* 170224 170323: gap of unknown length  
\* 170324 189235: contig of 18912 bp in length  
\* 189236 189335: gap of unknown length  
\* 189336 199422: contig of 10087 bp in length  
\* 199423 199522: gap of unknown length  
\* 199523 201282: contig of 1760 bp in length  
\* 201283 201382: gap of unknown length  
\* 201383 204316: contig of 2934 bp in length.

FEATURES  
source 1..204316  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="6"  
/map="6B2"  
/clone="RP23-62E15"

BASE COUNT 61234 a 39664 c 38680 g 61651 t 3087 others  
ORIGIN  
Query Match 63.6%; Score 645.4; DB 2: Length 204316;  
Best Local Similarity 80.5%; Pred. No. 3.2e-148;  
Matches 779; Conservative 0; Mismatches 187; Indels 2; Gaps 2;

QY	25	CATTAACTGATTTCCCTCTCTCATAGGAGATATGGGGACAATATPAATCCATCACAGAG	84
Db	117956	CATGAATTCAGTTTCCCTCTCTCATAGGAGATATGGGAACAATATGACATTATACACAGAG	118015
QY	85	TTCCCTCTACTGGGATTTCCCGGTGGCCCCAAGATTTCAGATGCTCCTTTGGGCTCTTC	144
Db	118016	TTCAATCTCTGNGATTTCTCTCTCAGACCCGAGAGATGCAGATGCTCCTTTGGCCCTCTTC	118075
QY	145	TCCCTGTCTACGCTCTTCACCCCTGCTGGGGAACGGGACCACTACTGGGGCTCATCTCACTG	204
Db	118076	TCCCTCTCTATGCTTCACCCCTGCTGGGGAATGGAACCACTGCTGGGGCTCATCTGCTTCG	118135
QY	205	GACTCCAGACTGCACGC-CCCCGTACTCTTCTCTCTCAACCTGGCGGCTCGTGACATC	263
Db	118136	GACTCCAGACTCCACACTCCCATGTACTTCTTCTCTCTCATCTGGCCATCGTTGACATT	118195
QY	264	GCCTACGCGCTGCAACACGCGTGCCCCGGATGCTGCTGTAACCTCTGCAATCCAGCCAAAGCC	323
Db	118196	GCCTATGCGCTGCAACACAGTGCCCCAGATGCTGCTGAACCTTTAGATCCAGTGAAGCCC	118255
QY	324	ATCTCCTTGGCGGCGCGCATGATGCAGACCTTCTGTTTCCACTTTTGCTGTACAGAA	383
Db	118256	ATTCCTATGCTGGGTGTATGACACAGACTTTCTCTTTTGACATTGCTATACAGAA	118315
QY	384	TGTCCTCTCCTGGTGTGATGTCCCTATGATCTGTACGTGCGCATCTGCCACCCCTCCGA	443
Db	118316	TGCTCTCTCCTAGTGTGATGTCCCTATGACCGGTATGTGCAATCTGTCACCCACTCCGT	118375
QY	444	TATTTGGCCATCATGACCTGGAGAGTCTGCATCACCCCTCGCGGTGACTCTCTGACCACT	503
Db	118376	TACTCTGCCATCATGAGCTGGAGAGTTTGACAGCACCAATGGCAGTGCATCTCTGATATT	118435
QY	504	GGAGTCCTTTATTCCTTGATTCATCTTGTGTACTTCTACCTTTAACCTTCTGTAGGCC	563
Db	118436	GGTGTCTGCTGTCTTTGATCCACCTCGTGTACTTTTACCTTACCTTCTGTGTGCT	118495
QY	564	CAGAAATTTATCACCTTTTGTGGAATCTTGCGCTGTCTCAAACTTGCCCTGTGCAGA	623
Db	118496	CAGAAAGTGAACCACTTTTCTGTGGAATTTACAGCTATCTCAAACTTGCCCTGTGCAGA	118554
QY	624	TACCACATCATGAGAGACATGCTCTGGCGGAGCAATTTCTGGGCTGTGGACCTT	683
Db	118555	CACACACCTCAATGAGAGAGTGTCTGCTGGGAGAGTGTCTGTCTGTGGGCCATT	118614
QY	684	GTCACAAATGTAGTTTCATATATATGTCATCTCTGTGTATCTTCAGATCCAAATCAAG	743
Db	118615	CTCCTCAATGTGTCTCTATATGATGATGATCTCTGTGTATCTTCAGATCCAAATCAAG	118674
QY	744	GGAAGTTCAGAGAAAGCCCTCTGACACCTGCTTCTCCACCTCTGTGTGATTGACTCTT	803
Db	118675	AGAGGGTCAAAAGAAAGCCCTTCTCACCCTGCTCTCCACCTCTGTGTGATTGATTGTT	118734
QY	804	TTATGGCAGCAGCATATATCATGATGATGTTGACCCAGATATGGGAACCCCAAGAGCAGAA	863
Db	118735	TTATGGTACCGCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	118794
QY	864	GAAATATCTCTGCTGTCTTACAGCCTCTTAAATCCATGCTCAATCCCTTATCTGTAG	923
Db	118795	GAAATATCTTGTGCTGTCTTACAGCCTCTTAAATCCATGCTCAATCCCTTATCTGTAG	118854
QY	924	TCTTAGGAACCTCAGAGTGAAGATACTTTGAAGAGAGTGTGGAGTAGAAAGGCTTT	983
Db	118855	TCTCAGGAATCTGACGTGAAGATATCATTTGAAAGAGTGTGAGGACACAGAGAGCTTT	118914
QY	984	ATGAAAG 991	
Db	118915	GTGACAAAG 118922	
RESULT 15			
LOCUS	AF399598	650 bp	DNA linear PRI 27-AUG-2001

DEFINITION	Homo sapiens clone OR2A7 olfactory receptor gene, partial cds.
ACCESSION	AF399598
VERSION	AF399598.1 GI:15293780
KEYWORDS	
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 650)
AUTHORS	Fuchs,T., Malecova,B., Linhart,C., Sharan,R., Khen,M., Herwig,R., Shmulevich,D., Elkon,R., Steinfath,M., O'Brien,J.K., Radelof,U., Lehrach,H., Olander,Z., Glusman,G., Lancet,D. and Shamir,R.
TITLE	DEFOG: A Practical Scheme for Deciphering Families of Genes
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 650)
AUTHORS	Fuchs,T., Malecova,B., Linhart,C., Sharan,R., Khen,M., Herwig,R., Shmulevich,D., Elkon,R., Steinfath,M., O'Brien,J.K., Radelof,U., Lehrach,H., Olander,Z., Glusman,G., Lancet,D. and Shamir,R.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Dept. of Molecular Genetics and the Crown Human Genome Center, Weizmann Institute of Science, Rehovot 76100, Israel
FEATURES	Location/Qualifiers
source	1..650
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="149.18104425"
	/clone="OR2A7"
	<1..>650
	/product="olfactory receptor"
	<1..>650
	/note="contains transmembrane regions 2-7"
CDS	/codon_start=1
	/product="olfactory receptor"
	/protein_id="AAK95083.1"
	/db_xref="GI:15293781"
	/translation="VVDIAYACNFVPRMLVNLHPAKPISFAGRMQTELFSTFAVTE CLLVVMSYDLVAICHPLRYLAIMRWVCITLAVTSWTGVLISLHLVLLPLPFC RPOKIHFECETLAVLKACADTHINENWLAGISGLVGPLSTIVSYMCILCALIQ IQSREVRKAFCTCFSHLCVIGLFYGTALIMYVGPRIYGNPKQKKYLPFHSLENP"
BASE COUNT	131 a 187 c 136 g 196 t
ORIGIN	
Query Match	62.9%; Score 637.4; DB 9; Length 650;
Best Local Similarity	99.7%; Pred. No. 2.3e-146;
Matches 649; Conservative	0; Mismatches 1; Indels 1; Gaps 1;
QY 252	GTCGTGCACATCGCCCTACCGCTGCACACACGCGTCCCGGATGCTGGTGAACCTCCAT 311
Db 1	GTCGTGCACATCGCCCTACCGCTGCACACACGCGTCCCGGATGCTGGTGAACCTCCAT 60
QY 312	CCAGCCCAAGCCCATCTCTTTGGCGGGCGGCGATGATGCACACCTTCTGTGTTTCCACTTTT 371
Db 61	CCAGCCCAAGCCCATCTCTTTGGCGGGCGGCGATGATGCACACCTTCTGTGTTTCCACTTTT 120
QY 372	GCTGTACAGAAATGTCTCTCTCTGTTGGTGTGATGTCTCTATGATCTGTACGTGGCCATCTGC 431
Db 121	GCTGTACAGAAATGTCTCTCTCTGTTGGTGTGATGTCTCTATGATCTGTACGTGGCCATCTGC 180
QY 432	CACCCCTCCGATATTTGGCCATCATGACCTGGAGAGCTCTGCATCACCCCTGGCGGTACT 491
Db 181	CACCCCTCCGATATTTGGCCATCATGACCTGGAGAGCTCTGCATCACCCCTGGCGGTACT 240
QY 492	TCCTGGACCACTGGAGTCTTTTATCCTTGATTCATCTGTGTACTTCTACCTTTACCC 551
Db 241	TCCTGGACCACTGGAGTCTTTTATCCTTGATTCATCTGTGTACTTCTACCTTTACCC 300
QY 552	TTCTGTAGGCCCCAGAAAATTTATCACTTTTTTTTGTGAATCTTGGCTGTTCTCAACT 611
Db 301	TTCTGTAGGCCCCAGAAAATTTATCAC-TTTTTTTGTGAATCTTGGCTGTTCTCAACT 359
QY 612	TGCGTGTGCAGATACCCATCATGAGACATGGTCTGGCCGGAGCAATTTCTGGGCT 671

```

Db 360 TGCCTGTGCAGATACCCACATCAATGAGAACATGCTCTTGCCGAGCAATTCTGGGCT 419
QY 672 GGTGGACCCCTGTCCACAATTGTAGTTTCATATATGTGCATCCCTGTGCTATCCCTCA 731
Db 420 GGTGGACCCCTGTCCACAATTGTAGTTTCATATATGTGCATCCCTGTGCTATCCCTCA 479
QY 732 GATCCAATCAAGGGAAGTTCAGAGAAAGCCTTGTGACACCTGCTTCGCCACCTCTGTGT 791
Db 480 GATCCAATCAAGGGAAGTTCAGAGAAAGCCTTGTGACACCTGCTTCGCCACCTCTGTGT 539
QY 792 GATGGACTCTTTATGSCACAGCCATTATPCANGTATGTGGACCCAGATATGGGAACCC 851
Db 540 GATGGACTCTTTATGSCACAGCCATTATPCANGTATGTGGACCCAGATATGGGAACCC 599
QY 852 CAAGGAGCAGAGAATAATCTCTGCTGTTCACAGCCCTTTAATCCCAT 902
Db 600 CAAGGAGCAGAGAATAATCTCTGCTGTTCACAGCCCTTTAATCCCAT 650

```

Search completed: May 2, 2003, 05:55:22  
 Job time : 3242 secs

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 2, 2003, 04:56:35 ; Search time 1671 Seconds  
(without alignments)  
717.034 Million cell updates/sec

Title: US-09-898-586-23

Perfect score: 1014  
Sequence: 1 taacactctcctaaca.....tatgcatgtgactgacaa 1014

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1014	100.0	1014	9 US-09-898-586-23	Sequence 23, Appl
2	1014	100.0	1014	10 US-09-761-288-23	Sequence 23, Appl
3	987.4	97.4	1012	9 US-09-898-586-21	Sequence 21, Appl
4	987.4	97.4	1012	10 US-09-761-288-21	Sequence 21, Appl
5	983.6	97.0	1040	9 US-09-898-586-3	Sequence 3, Appl
6	983.6	97.0	1040	9 US-09-898-586-34	Sequence 34, Appl
7	983.6	97.0	1040	10 US-09-761-288-3	Sequence 3, Appl
8	983.6	97.0	1040	10 US-09-761-288-34	Sequence 34, Appl
9	545	53.7	1957	10 US-09-864-761-13963	Sequence 13963, A
10	538.2	53.1	2282	9 US-10-098-841-92	Sequence 92, Appl
11	538.2	53.1	2282	10 US-09-747-835A-62	Sequence 62, Appl
12	536.6	52.9	1788	10 US-09-747-835A-35	Sequence 35, Appl
13	536.6	52.9	2735	10 US-09-747-835A-34	Sequence 34, Appl
14	535.2	52.8	933	10 US-09-864-761-30527	Sequence 30527, A
15	511.4	50.4	1782	10 US-09-747-835A-41	Sequence 41, Appl
16	502.4	49.5	927	10 US-09-747-835A-43	Sequence 43, Appl
17	481.2	47.5	1713	9 US-10-024-494-1	Sequence 1, Appl
18	432	42.6	456	10 US-09-812-102-72	Sequence 72, Appl
19	414	40.8	933	10 US-09-886-055-486	Sequence 486, App

20	402.2	39.7	930	10 US-09-886-055-472	Sequence 472, App
21	390	38.5	933	10 US-09-886-055-484	Sequence 484, App
22	359	35.4	936	10 US-09-886-055-482	Sequence 482, App
23	273.2	26.9	400	9 US-09-918-995-16242	Sequence 16242, A
24	257.2	25.4	585	10 US-09-747-835A-37	Sequence 37, Appl
25	250.2	24.7	930	10 US-09-771-730-111	Sequence 111, App
26	250.2	24.7	1038	10 US-09-886-055-322	Sequence 322, App
27	249.4	24.6	921	10 US-09-771-730-135	Sequence 135, App
28	246.4	24.3	1041	10 US-09-886-055-228	Sequence 228, App
29	245	24.2	975	10 US-09-886-055-326	Sequence 326, App
30	242.6	23.9	936	10 US-09-886-055-10	Sequence 10, Appl
31	241.6	23.8	948	10 US-09-886-055-176	Sequence 176, App
32	240.2	23.7	1954	10 US-09-864-761-5098	Sequence 5098, Ap
33	239.2	23.6	1031	9 US-09-974-591-5	Sequence 5, Appl
34	239.2	23.6	1031	9 US-09-974-591-21	Sequence 21, Appl
35	238.8	23.6	916	9 US-09-974-591-15	Sequence 15, Appl
36	237.6	23.4	939	10 US-09-886-055-290	Sequence 290, App
37	237.6	23.4	945	10 US-09-886-055-378	Sequence 378, App
38	236.4	23.3	960	10 US-09-886-055-70	Sequence 70, Appl
39	235.8	23.3	969	10 US-09-886-055-254	Sequence 254, App
40	235.4	23.2	936	10 US-09-886-055-470	Sequence 470, App
41	235	23.2	939	10 US-09-886-055-66	Sequence 66, Appl
42	235	23.2	1050	9 US-09-974-591-1	Sequence 1, Appl
43	235	23.2	1050	9 US-09-974-591-3	Sequence 3, Appl
44	235	23.2	1050	9 US-09-974-591-7	Sequence 7, Appl
45	234.6	23.1	942	10 US-09-886-055-6	Sequence 6, Appl

## ALIGNMENTS

RESULT 1  
US-09-898-586-23  
; Sequence 23, Application US/09898586  
; General Information: US2003007794A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: MacDougall, John R  
; APPLICANT: Smithson, Glenda  
; TITLE OF INVENTION: No. US2003007794A1el Polypeptides and Nucleic Acids Encoding  
; FILE REFERENCE: 15966-638CIP  
; CURRENT APPLICATION NUMBER: US/09/898, 586  
; CURRENT FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 60/177, 839  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 60/176, 134  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 60/175, 989  
; PRIOR FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: 60/218, 324  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220, 253  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: 60/178, 191  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/178, 227  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/220, 590  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: 09/761, 288  
; PRIOR FILING DATE: 2001-01-16  
; NUMBER OF SEQ ID NOS: 104  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 1014  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-898-586-23

Query Match 100.0%; Score 1014; DB 9; length 1014;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TAAACACTTCTCTTAACCATGAGCATTAACCTGATTTCTCTGTCATAGGATATGGGG 60  
 Db 1 TAAACACTTCTCTTAACCATGAGCATTAACCTGATTTCTCTGTCATAGGATATGGGG 60  
 OY 61 GACAATATATACATCCATCACAGAGTTCCTCTACTGGATTTCCCGTGGCCCAAGATT 120  
 Db 61 GACAATATATACATCCATCACAGAGTTCCTCTACTGGATTTCCCGTGGCCCAAGATT 120  
 OY 121 CAGATGCTCTCTTTGGGCTCTTCCCTGTTCTACGTCCTCAACCTGCTGGGAACGGG 180  
 Db 121 CAGATGCTCTCTTTGGGCTCTTCCCTGTTCTACGTCCTCAACCTGCTGGGAACGGG 180  
 OY 181 ACCATAGTGGGCTCATCTCACTGAGCTCAGACTGCAGCCGCCCTGACTCTCTCTCT 240  
 Db 181 ACCATAGTGGGCTCATCTCACTGAGCTCAGACTGCAGCCGCCCTGACTCTCTCTCT 240  
 OY 241 CACACCTGGCGGTGTCGACATGCGCTGACGCTGCAACACGGGTGCCCCGGATGCTGTGA 300  
 Db 241 CACACCTGGCGGTGTCGACATGCGCTGACGCTGCAACACGGGTGCCCCGGATGCTGTGA 300  
 OY 301 ACCTGCTGATCCAGCCCAAGCCCATCTCTTTGGGGCCGATGATGAGACCTTCTGT 360  
 Db 301 ACCTGCTGATCCAGCCCAAGCCCATCTCTTTGGGGCCGATGATGAGACCTTCTGT 360  
 OY 361 TTTCACACTTTGCTGTACAGAAATGTCTCTCTGCTGCTGCTGCTATGATCTGTAGC 420  
 Db 361 TTTCACACTTTGCTGTACAGAAATGTCTCTCTGCTGCTGCTGCTATGATCTGTAGC 420  
 OY 421 TGGCCATCTGCGACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCACCC 480  
 Db 421 TGGCCATCTGCGACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCACCC 480  
 OY 481 TCGCGGTGACTTCTCTGAGCCACTGAGTCTTTTATCTTATCTGATTCATCTGTCTACTTC 540  
 Db 481 TCGCGGTGACTTCTCTGAGCCACTGAGTCTTTTATCTTATCTGATTCATCTGTCTACTTC 540  
 OY 541 TACCTTTACCTTCTGTAAGGCCCCAGAAAAATTATCACTTTTGTGAAATCTTGGCT 600  
 Db 541 TACCTTTACCTTCTGTAAGGCCCCAGAAAAATTATCACTTTTGTGAAATCTTGGCT 600  
 OY 601 GTTCCTCAAACTTGCCTGTGAGATATCCACATCATGAGAACATGCTTGGCCGAGCA 660  
 Db 601 GTTCCTCAAACTTGCCTGTGAGATATCCACATCATGAGAACATGCTTGGCCGAGCA 660  
 OY 661 ATTCTGGGCTGTGGGACCTTGTCCACAAATTTAGTTTATATATGTCATCTCTGT 720  
 Db 661 ATTCTGGGCTGTGGGACCTTGTCCACAAATTTAGTTTATATATGTCATCTCTGT 720  
 OY 721 GCTATCTCTGAGATCCATCAAGGGAAGTTCAGAGAAAGCCTTCTGCACCTCTCTCC 780  
 Db 721 GCTATCTCTGAGATCCATCAAGGGAAGTTCAGAGAAAGCCTTCTGCACCTCTCTCC 780  
 OY 781 CACCTCTGTGATTTGAGTCTTTTATGGCACAGCCATTTATCATGTATGTGGACCCAGA 840  
 Db 781 CACCTCTGTGATTTGAGTCTTTTATGGCACAGCCATTTATCATGTATGTGGACCCAGA 840  
 OY 841 TATGGGAACCCCAAGAGAGCAAGAATATCTCTCTGCTTTTACAGCCCTCTTAATCCC 900  
 Db 841 TATGGGAACCCCAAGAGAGCAAGAATATCTCTCTGCTTTTACAGCCCTCTTAATCCC 900  
 OY 901 ATGCTCAATCCCTTATCTGTAGTCTTAGGAAGTCAAGTGAAGATATCTTTGAAGAGA 960  
 Db 901 ATGCTCAATCCCTTATCTGTAGTCTTAGGAAGTCAAGTGAAGATATCTTTGAAGAGA 960  
 OY 961 GTGCTGGAGTAGAAGGGCTTTATGAAAAGGATTTATGCAATTTGACTGACAA 1014  
 Db 961 GTGCTGGAGTAGAAGGGCTTTATGAAAAGGATTTATGCAATTTGACTGACAA 1014

RESULT 2  
 US-09-761-288-23  
 ; Sequence 23, Application US/09761288  
 ; Patent No. US20020065405A1

GENERAL INFORMATION:  
 ; APPLICANT: Padigaru, Muralidhara  
 ; APPLICANT: Prayaga, Suchirdas  
 ; APPLICANT: Prayager, Raymond J  
 ; APPLICANT: Mishra, Vishnu  
 ; APPLICANT: Tchernev, Velizar  
 ; APPLICANT: Spytek, Kimberly  
 ; APPLICANT: Li, Li  
 ; TITLE OF INVENTION: NO. US20020065405A1 Polypeptides and Nucleic Acids Encoding  
 ; FILE REFERENCE: 15966-638  
 ; CURRENT APPLICATION NUMBER: 2001-01-16  
 ; CURRENT FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: 60/177,839  
 ; PRIOR FILING DATE: 2000-01-25  
 ; PRIOR APPLICATION NUMBER: 60/176,134  
 ; PRIOR FILING DATE: 2000-01-14  
 ; PRIOR APPLICATION NUMBER: 60/175,989  
 ; PRIOR FILING DATE: 2000-01-13  
 ; PRIOR APPLICATION NUMBER: 60/218,324  
 ; PRIOR FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: 60/220,253  
 ; PRIOR FILING DATE: 2000-07-24  
 ; PRIOR APPLICATION NUMBER: 60/178,191  
 ; PRIOR FILING DATE: 2000-01-26  
 ; PRIOR APPLICATION NUMBER: 60/178,227  
 ; PRIOR FILING DATE: 2000-01-26  
 ; PRIOR APPLICATION NUMBER: 60/220,590  
 ; PRIOR FILING DATE: 2000-07-25  
 ; NUMBER OF SEQ ID NOS: 95  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 23  
 ; LENGTH: 1014  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-761-288-23

Query Match 100.0%; Score 1014; DB 10; Length 1014;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 TAAACACTTCTCTTAACCATGAGCATTAACCTGATTTCTCTGTCATAGGATATGGGG 60  
 Db 1 TAAACACTTCTCTTAACCATGAGCATTAACCTGATTTCTCTGTCATAGGATATGGGG 60  
 OY 61 GACAATATATACATCCATCACAGAGTTCCTCTACTGGATTTCCCGTGGCCCAAGATT 120  
 Db 61 GACAATATATACATCCATCACAGAGTTCCTCTACTGGATTTCCCGTGGCCCAAGATT 120  
 OY 121 CAGATGCTCTCTTTGGGCTCTTCCCTGTTCTACGTCCTCAACCTGCTGGGAACGGG 180  
 Db 121 CAGATGCTCTCTTTGGGCTCTTCCCTGTTCTACGTCCTCAACCTGCTGGGAACGGG 180  
 OY 181 ACCATAGTGGGCTCATCTCACTGAGCTCAGACTGCAGCCGCCCTGACTCTCTCTCT 240  
 Db 181 ACCATAGTGGGCTCATCTCACTGAGCTCAGACTGCAGCCGCCCTGACTCTCTCTCT 240  
 OY 241 CACACCTGGCGGTGTCGACATGCGCTGACGCTGCAACACGGGTGCCCCGGATGCTGTGA 300  
 Db 241 CACACCTGGCGGTGTCGACATGCGCTGACGCTGCAACACGGGTGCCCCGGATGCTGTGA 300  
 OY 301 ACCTGCTGATCCAGCCCAAGCCCATCTCTTTGGGGCCGATGATGAGACCTTCTGT 360  
 Db 301 ACCTGCTGATCCAGCCCAAGCCCATCTCTTTGGGGCCGATGATGAGACCTTCTGT 360  
 OY 361 TTTCACACTTTGCTGTACAGAAATGTCTCTCTGCTGCTGCTATGATCTGTAGC 420  
 Db 361 TTTCACACTTTGCTGTACAGAAATGTCTCTCTGCTGCTGCTATGATCTGTAGC 420  
 OY 421 TGGCCATCTGCGACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCACCC 480  
 Db 421 TGGCCATCTGCGACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCACCC 480  
 OY 481 TCGCGGTGACTTCTCTGAGCCACTGAGTCTTTTATCTTATCTGATTCATCTGTCTACTTC 540



```
|||||
Db 481 TCGGGGTGACTTCTGGAGCACTGGAGTCTTTTATCTTGATCATCTGTGTACTTC 540
QY 541 TACCTTACCCTTCTGTAGGCCCCAGAAATTTATCACTTTTGTGAATCTGGCT 600
Db 541 TACCTTACCCTTCTGTAGGCCCCAGAAATTTATCACTTTTGTGAATCTGGCT 600
QY 601 GTTCTCAACTGGCTGTGCAGATACCCACATCATGAGAACATGGCTTGGCCGAGCA 660
Db 601 GTTCTCAACTGGCTGTGCAGATACCCACATCATGAGAACATGGCTTGGCCGAGCA 660
QY 661 ATTCTGGGCTGGTGGGACCCTTGTCCACAAATGTAGTTCAATATGTGCATCTGT 720
Db 661 ATTCTGGGCTGGTGGGACCCTTGTCCACAAATGTAGTTCAATATGTGCATCTGT 720
QY 721 GCTATCCTTCAGATCCAAATCAAGGGAAGTTCAGAGAAAGCCTTCTGCACCTGTCTCC 780
Db 721 GCTATCCTTCAGATCCAAATCAAGGGAAGTTCAGAGAAAGCCTTCTGCACCTGTCTCC 780
QY 781 CACCTCTGTGTGATGGACTCTTTATGGACAGCCATTATCATGTATGTGACCAGA 840
Db 781 CACCTCTGTGTGATGGACTCTTTATGGACAGCCATTATCATGTATGTGACCAGA 840
QY 841 TATGGGAACCCCAAGAGCAGAGAATAATATCTCTGCTTTCACAGCCTCTTAAATCC 900
Db 841 TATGGGAACCCCAAGAGCAGAGAATAATATCTCTGCTTTCACAGCCTCTTAAATCC 900
QY 901 ATGCTCAATCCCTTATCTGTAGTCTTAGAAGTTCAGAGTGAAGATACTTGAAGAGA 960
Db 901 ATGCTCAATCCCTTATCTGTAGTCTTAGAAGTTCAGAGTGAAGATACTTGAAGAGA 960
QY 961 GTGCTGGAGTAGAAGGCTTTATGAAAGGATTATGGCATTTGACTGACAA 1014
Db 961 GTGCTGGAGTAGAAGGCTTTATGAAAGGATTATGGCATTTGACTGACAA 1014
```

## RESULT 3

US-09-898-586-21

; Sequence 21, Application US/09898586

; Publication No. US20030077794A1

; GENERAL INFORMATION:

; APPLICANT: Gerlach, Valerie L.

; APPLICANT: MacDougall, John R

; APPLICANT: Smithson, Glenda

; TITLE OF INVENTION: No. US20030077794A1 Polypeptides and Nucleic Acids Encoding Sam

; FILE REFERENCE: 15966-638CIP

; CURRENT APPLICATION NUMBER: US/09/898,586

; PRIOR APPLICATION NUMBER: 60/177,839

; PRIOR FILING DATE: 2000-01-25

; PRIOR APPLICATION NUMBER: 60/176,134

; PRIOR FILING DATE: 2000-01-14

; PRIOR APPLICATION NUMBER: 60/175,989

; PRIOR FILING DATE: 2000-01-13

; PRIOR APPLICATION NUMBER: 60/218,324

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/220,253

; PRIOR FILING DATE: 2000-07-24

; PRIOR APPLICATION NUMBER: 60/178,191

; PRIOR FILING DATE: 2000-01-26

; PRIOR APPLICATION NUMBER: 60/178,227

; PRIOR FILING DATE: 2000-01-26

; PRIOR APPLICATION NUMBER: 60/220,590

; PRIOR FILING DATE: 2000-07-25

; PRIOR APPLICATION NUMBER: 09/761,288

; PRIOR FILING DATE: 2001-01-16

; NUMBER OF SEQ ID NOS: 104

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 21

; LENGTH: 1012

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-898-586-21

Query Match 97.4%; Score 987.4; DB 9; Length 1012;  
Best Local Similarity 99.7%; Pred. No. 5.2e-312;  
Matches 1010; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
QY 2 AAACACTTCTCTAAACCATGAGCATTTAACTTGATTTCTCTGTCTATAGGATATGGGG 61
Db 1 AAACACTTCTCTAAACCATGAGCATTTAACTTGATTTCTCTGTCTATAGGATATGGAG 60
QY 62 ACAATATACATCCATCACAGAGTTCCTCTACTGGGATTTCCCGTGGCCCAAGATTC 121
Db 61 ACAATATACATCCATCACAGAGTTCCTCTACTGGGATTTCCCGTGGCCCAAGATTC 120
QY 122 AGATGCTCTCTTTGGGCTCTTCTCCTGTCTACGTCTTCAACCTGCTGGGGAAGGGA 181
Db 121 AGATGCTCTCTTTGGGCTCTTCTCCTGTCTACGTCTTCAACCTGCTGGGGAAGGGA 180
QY 182 CCATACCTGGGCTATCTCACATGACCTCCAGACTGCACGCCCC-TGTACTTCTCTCT 240
Db 181 CCATACCTGGGCTATCTCACATGACCTCCAGACTGCACGCCCCATGTACTTCTCTCT 240
QY 241 CACACCTGGGCTGTGCACATGCGCTTACGCTGCAACACAGGTGCCCCGATGCTGTGA 300
Db 241 CACACCTGGGCTGTGCACATGCGCTTACGCTGCAACACAGGTGCCCCGATGCTGTGA 300
QY 301 ACCCTCTGATCCAGCCCAAGCCCATCTCTTTGCGGGCCGATGAGACAGACCTTCTGT 360
Db 301 ACCCTCTGATCCAGCCCAAGCCCATCTCTTTGCGGGCCGATGAGACAGACCTTCTGT 360
QY 361 TTTCACCTTTTGTCTGTACAGAAATGTCTCTCTGCTGTGTGTATGATGATGATGATG 420
Db 361 TTTCACCTTTTGTCTGTACAGAAATGTCTCTCTGCTGTGTGTATGATGATGATGATG 420
QY 421 TGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCACCC 480
Db 421 TGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCACCC 480
QY 481 TCGGGGTGACTTCTCTGGAGCACTGGAGTCTTTATCCTTGAATCTGTGTTACTTC 540
Db 481 TCGGGGTGACTTCTCTGGAGCACTGGAGTCTTTATCCTTGAATCTGTGTTACTTC 540
QY 541 TACCTTTACCCCTCTGTAGGCCCCAGAAATTTATCACTTTTGTGAATCTGTGCT 600
Db 541 TACCTTTACCCCTCTGTAGGCCCCAGAAATTTATCAC-TTTTGTGTGAATCTGTGCT 599
QY 601 GTTCTCAAACTTGGCTGTGCAGATACCCACATCATGAGAACATGGTCTTGGCCGAGCA 660
Db 600 GTTCTCAAACTTGGCTGTGCAGATACCCACATCATGAGAACATGGTCTTGGCCGAGCA 659
QY 661 ATTCTGGGCTGTGGGACCCTTGTCCACAAATGTAGTTCAATATGTGCATCTCTGT 720
Db 660 ATTCTGGGCTGTGGGACCCTTGTCCACAAATGTAGTTCAATATGTGCATCTCTGT 719
QY 721 GCTATCCTTCAGATCCAAATCAAGGGAAGTTCAGAGAAAGCCTTCTGCACCTGTCTCC 780
Db 720 GCTATCCTTCAGATCCAAATCAAGGGAAGTTCAGAGAAAGCCTTCTGCACCTGTCTCC 779
QY 781 CACCTCTGTGTGATGGACTCTTTATGGCACAGCCATTATCATGTATGTGAGCCAGA 840
Db 780 CACCTCTGTGTGATGGACTCTTTATGGCACAGCCATTATCATGTATGTGAGCCAGA 839
QY 841 TATGGGAACCCCAAGAGCAGAGAATAATATCTCTGCTGTTCACAGCCTCTTAAATCC 900
Db 840 TATGGGAACCCCAAGAGCAGAGAATAATATCTCTGCTGTTCACAGCCTCTTAAATCC 899
QY 901 ATGCTCAATCCCTTATCTGTAGTCTTAGAAGTTCAGAGTGAAGATACTTGAAGAGA 960
Db 900 ATGCTCAATCCCTTATCTGTAGTCTTAGAAGTTCAGAGTGAAGATACTTGAAGAGA 959
QY 961 GTGCTGGAGTAGAAGGCTTTATGAAAGGATTATGGCATTTGACTGACAA 1013
Db 960 GTGCTGGAGTAGAAGGCTTTATGAAAGGATTATGGCATTTGACTGACAA 1012
```

## RESULT 4

US-09-761-288-21  
; Sequence 21, Application US/09761288  
; Patent No. US20020065405A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Prayaga, Sudhirdas  
; APPLICANT: Taupier, Raymond J  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spylek, Kimberky  
; APPLICANT: Li, Li  
; TITLE OF INVENTION: No. US20020065405A1el Polypeptides and Nucleic Acids Encoding Sam  
; FILE REFERENCE: 15966-638  
; CURRENT APPLICATION NUMBER: US/09/761,288  
; CURRENT FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/177,839  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 60/176,134  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 60/175,989  
; PRIOR FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: 60/218,324  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220,253  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: 60/178,191  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/178,227  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/220,590  
; PRIOR FILING DATE: 2000-07-25  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 1012  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-761-288-21

Query Match 97.4%; Score 987.4; DB 10; Length 1012;

Best Local Similarity 99.7%; Pred. No. 5.2e-312;

Matches 1010; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```
QY 2 AAACACTTCTCTAACCATGAGCATTAATGATTTCTCTGTGATAGGGAATGGGG 61
   |||||||
Db 1 AAACACTTCTCTAACCATGAGCATTAATGATTTCTCTGTGATAGGGAATGGAG 60

QY 62 ACAATATACATCCATCAGAGTCTCTCTACTGGGATTTCCGTTGGCCCAAGATTG 121
   |||||||
Db 61 ACAATATACATCCATCAGAGTCTCTCTACTGGGATTTCCGTTGGCCCAAGATTG 120

QY 122 AGATGCTCTCTTTGGGCTCTTCTCCCTGTTCAGTCTTCAACCCTGTGGGGAACGGGA 181
   |||||||
Db 121 AGATGCTCTCTTTGGGCTCTTCTCCCTGTTCAGTCTTCAACCCTGTGGGGAACGGGA 180

QY 182 CCATACCTGGGCTCATCTGACTGACTCCAGACTGACGCCCCC-TGTACTTCTCTCT 240
   |||||||
Db 181 CCATACCTGGGCTCATCTGACTGACTCCAGACTGACGCCCCCATGTACTTCTCTCT 240

QY 241 CACACCTGGGCTGCTGACATCGCCTAGGCTGCAACAGGTTGCCCGGATGCTGTGA 300
   |||||||
Db 241 CACACCTGGGCTGCTGACATCGCCTAGGCTGCAACAGGTTGCCCGGATGCTGTGA 300

QY 301 ACCTCTGCATCCAGCCCATCTCTTTGGCGGCGGCATGATGACAGACTTTCTGT 360
   |||||||
Db 301 ACCTCTGCATCCAGCCCATCTCTTTGGCGGCGGCATGATGACAGACTTTCTGT 360

QY 361 TTTCACCTTTTGTCTGACAGATGTCTCTCTGTGTGATGTCTATGATCTGTAGG 420
   |||||||
Db 361 TTTCACCTTTTGTCTGACAGATGTCTCTCTGTGTGATGTCTATGATCTGTAGG 420
```

```
QY 421 TGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCACCC 480
   |||||||
Db 421 TGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCACCC 480

QY 481 TCGCGGTGACTCTCTGGACCACTGAGAGTCTTTATCTTGATTCATCTGTGTTACTTC 540
   |||||||
Db 481 TCGCGGTGACTCTCTGGACCACTGAGAGTCTTTATCTTGATTCATCTGTGTTACTTC 540

QY 541 TACCTTACCTCTCTGTAGGCCCCAGAAAATTATCACTTTTGTGAAATCTTGCT 600
   |||||||
Db 541 TACCTTACCTCTCTGTAGGCCCCAGAAAATTATCAC-TTTTGTGAAATCTTGCT 599

QY 601 GTTCTCAACTTGCCTGTGACAGATACCCACATCATGAGAACATGCTTTGGCCGAGCA 660
   |||||||
Db 600 GTTCTCAACTTGCCTGTGACAGATACCCACATCATGAGAACATGCTTTGGCCGAGCA 659

QY 661 ATTCTTGGGCTGTGGACCCCTGTGTCACAAATTTAGTTTCAATATATGTGCAATCCTCTGT 720
   |||||||
Db 660 ATTCTTGGGCTGTGGACCCCTGTGTCACAAATTTAGTTTCAATATATGTGCAATCCTCTGT 719

QY 721 GGTATCCCTTCAGATCCAAATCAAGGGAAGTTTCAGAGAAAGCCTTCTGCACCTCTCTCC 780
   |||||||
Db 720 GGTATCCCTTCAGATCCAAATCAAGGGAAGTTTCAGAGAAAGCCTTCTGCACCTCTCTCC 779

QY 781 CACCTCTGTGATTTGAGCTCTTTATGCGACAGAGCATTTATCATGTATGTGACCCAGA 840
   |||||||
Db 780 CACCTCTGTGATTTGAGCTCTTTATGCGACAGAGCATTTATCATGTATGTGACCCAGA 839

QY 841 TATGGGAACCCCAAGAGCAGAGAATAATATCTCTGCTGTTTCACAGCCTCTTAAATCC 900
   |||||||
Db 840 TATGGGAACCCCAAGAGCAGAGAATAATATCTCTGCTGTTTCACAGCCTCTTAAATCC 899

QY 901 ATGCTCAATCCCTTATCTGTAGTCTTAGAAGTCTGAGAAATAGTTGAAGAGA 960
   |||||||
Db 900 ATGCTCAATCCCTTATCTGTAGTCTTAGAAGTCTGAGAAATAGTTGAAGAGA 959

QY 961 GTGCTGGAGTAGAAGGCTTTATGAAAAGGATTAAGCATTTGACTGACA 1013
   |||||||
Db 960 GTGCTGGAGTAGAAGGCTTTATGAAAAGGATTAAGCATTTGACTGACA 1012
```

## RESULT 5

US-09-898-586-3

; Sequence 3, Application US/09898586

; Publication No. US20030077794A1

; GENERAL INFORMATION:

; APPLICANT: Gerlach, Valerie L

; APPLICANT: MacDougall, John R

; APPLICANT: Smithson, Glenda

; TITLE OF INVENTION: No. US20030077794A1el Polypeptides and Nucleic Acids Encoding

; FILE REFERENCE: 15966-638CIP

; CURRENT APPLICATION NUMBER: US/09/898,586

; PRIOR APPLICATION NUMBER: 60/177,839

; PRIOR FILING DATE: 2000-01-25

; PRIOR APPLICATION NUMBER: 60/176,134

; PRIOR FILING DATE: 2000-01-14

; PRIOR APPLICATION NUMBER: 60/175,989

; PRIOR FILING DATE: 2000-01-13

; PRIOR APPLICATION NUMBER: 60/218,324

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/220,253

; PRIOR FILING DATE: 2000-07-24

; PRIOR APPLICATION NUMBER: 60/178,191

; PRIOR FILING DATE: 2000-01-26

; PRIOR APPLICATION NUMBER: 60/178,227

; PRIOR FILING DATE: 2000-01-26

; PRIOR APPLICATION NUMBER: 60/220,590

; PRIOR FILING DATE: 2000-07-25

; PRIOR APPLICATION NUMBER: 09/761,288

; PRIOR FILING DATE: 2001-01-16

; NUMBER OF SEQ ID NOS: 104

; SOFTWARE: Patentln Ver. 2.1

; SEQ ID NO 3  
; LENGTH: 1040  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-898-586-3

Query Match 97.0%; Score 983.6; DB 9; Length 1040;  
Best Local Similarity 99.4%; Pred. No. 9.2e-311;  
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

OY 1 TAACACTTCTCTTAACCATGAGCATTAACCTGATTCCTCTGTCATAGGATATGGG 60  
DB 28 TAACACTTCTCTTAACCATGAGCATTAACCTGATTCCTCTGTCATAGGATATGGG 87  
OY 61 GACATATATACATCCATCAGAGATTCCTCTACTGGGATTTCCCGTTGGCCCAAGCAT 120  
DB 88 GACATATATACATCCATCAGAGATTCCTCTACTGGGATTTCCCGTTGGCCCAAGCAT 147  
OY 121 CAGATGCTCTCTTTGGGCTCTTCCCTGTCTACGCTTTCAACCTGCTGGGGAACGG 180  
DB 148 CAGATGCTCTCTTTGGGCTCTTCCCTGTCTACGCTTTCAACCTGCTGGGGAACGG 207  
OY 181 ACCATAGTGGGCTCATCTCACTGGAGCTCCAGACTGCAGCGCCCTGTTACTTCTCTC 239  
DB 208 ACCATAGTGGGCTCATCTCACTGGAGCTCCAGACTGCAGCGCCCTGTTACTTCTCTC 267  
OY 240 TCACACCTGGGGGTCTGTCACATCGCTTACGCTGCAACAGGGTGCCCGGATGCTGTG 299  
DB 268 TCACACCTGGGGGTCTGTCACATCGCTTACGCTGCAACAGGGTGCCCGGATGCTGTG 327  
OY 300 AACCTCTGCATCCAGCCAGCCCATCTCTTGGCGGGCATGATGACAGACCTTTCTG 359  
DB 328 AACCTCTGCATCCAGCCAGCCCATCTCTTGGCGGGCATGATGACAGACCTTTCTG 387  
OY 360 TTTTCCACTTTTGTCTGTCACAGATGTCCTCTGTTGGTGTATGTCATGATCTGTAC 419  
DB 388 TTTTCCACTTTTGTCTGTCACAGATGTCCTCTGTTGGTGTATGTCATGATCTGTAC 447  
OY 420 GTGGCCATCTGGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCTGCATCAC 479  
DB 448 GTGGCCATCTGGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCTGCATCAC 507  
OY 480 CTGGGGTGAATCTCTGGAGCAGTGTCTTTATCCCTGATTCATCTGTGTACTT 539  
DB 508 CTGGGGTGAATCTCTGGAGCAGTGTCTTTATCCCTGATTCATCTGTGTACTT 567  
OY 540 CTACCTTTACCTCTCTGTAGGCCCCAGAAATTTATCATCTTTTGTGAATCTGGC 599  
DB 568 CTACCTTTACCTCTCTGTAGGCCCCAGAAATTTATCATCTTTTGTGAATCTGGC 626  
OY 600 TGTCTCAAACTTGCTGTGAGATACCCCATCAATGAGACATGCTTGGCCGAGC 659  
DB 627 TGTCTCAAACTTGCTGTGAGATACCCCATCAATGAGACATGCTTGGCCGAGC 686  
OY 660 AATTCTGGGCTGTGGAGCCCTTGTCCCAATTTGATTTCAATATATGTGCATCTCTG 719  
DB 687 AATTCTGGGCTGTGGAGCCCTTGTCCCAATTTGATTTCAATATATGTGCATCTCTG 746  
OY 720 TGTATCTCTGAGATCCATCAAGGAGTTCAGAGGAAAGCCTTCTGCACCTGCTCTC 779  
DB 747 TGTATCTCTGAGATCCATCAAGGAGTTCAGAGGAAAGCCTTCTGCACCTGCTCTC 806  
OY 780 CCACCTCTGTGTGATTTGAGCTCTTTATGGCACAGCCATATCATGTATGTGGACCCAG 839  
DB 807 CCACCTCTGTGTGATTTGAGCTCTTTATGGCACAGCCATATCATGTATGTGGACCCAG 866  
OY 840 ATATGGGAACCCCAAGAGAGAGAAATATCTCTGCTGTTTCAAGCCTCTTTAATCC 899  
DB 867 ATATGGGAACCCCAAGAGAGAGAAATATCTCTGCTGTTTCAAGCCTCTTTAATCC 926  
OY 900 CATGCTCAATCCCTTATCTGTAGTCTTGAAGAACTCAGAAAGTGAAGATATCTTTGAAG 959  
DB 927 CATGCTCAATCCCTTATCTGTAGTCTTGAAGAACTCAGAAAGTGAAGATATCTTTGAAG 986

OY 960 AGTCTGGGAGTAGAAGGGCTTTATGAAAAGATTAAGCATTTGTACTGACA 1013  
DB 987 AGTCTGGGAGTAGAAGGGCTTTATGAAAAGATTAAGCATTTGTACTGACA 1040

RESULT 6  
US-09-898-586-34

; Sequence 34, Application US/09898586  
; Publication No. US2003007794A1  
; GENERAL INFORMATION:  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: MacDougall, John R  
; APPLICANT: Smithson, Glenda  
; TITLE OF INVENTION: No. US2003007794A1 Polypeptides and Nucleic Acids Encoding  
; FILE REFERENCE: 15966-638CIP  
; CURRENT APPLICATION NUMBER: US/09/898,586  
; PRIOR FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 60/177,839  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 60/176,134  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 60/175,989  
; PRIOR FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: 60/218,324  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220,253  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: 60/178,191  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/178,227  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/220,590  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: 09/761,288  
; PRIOR FILING DATE: 2001-01-16  
; NUMBER OF SEQ ID NOS: 104  
; SOFTWARE: Patentlin Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 1040  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-898-586-34

Query Match 97.0%; Score 983.6; DB 9; Length 1040;  
Best Local Similarity 99.4%; Pred. No. 9.2e-311;  
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

OY 1 TAACACTTCTCTTAACCATGAGCATTAACCTGATTCCTCTGTCATAGGATATGGG 60  
DB 28 TAACACTTCTCTTAACCATGAGCATTAACCTGATTCCTCTGTCATAGGATATGGG 87  
OY 61 GACATATATACATCCATCAGAGATTCCTCTACTGGGATTTCCCGTTGGCCCAAGCAT 120  
DB 88 GACATATATACATCCATCAGAGATTCCTCTACTGGGATTTCCCGTTGGCCCAAGCAT 147  
OY 121 CAGATGCTCTCTTTGGGCTCTTCCCTGTCTACGCTTTCAACCTGCTGGGGAACGG 180  
DB 148 CAGATGCTCTCTTTGGGCTCTTCCCTGTCTACGCTTTCAACCTGCTGGGGAACGG 207  
OY 181 ACCATAGTGGGCTCATCTCACTGGAGCTCCAGACTGCAGCGCCCTGTTACTTCTCTC 239  
DB 208 ACCATAGTGGGCTCATCTCACTGGAGCTCCAGACTGCAGCGCCCTGTTACTTCTCTC 267  
OY 240 TCACACCTGGGGGTCTGTCACATCGCTTACGCTGCAACAGGGTGCCCGGATGCTGTG 299  
DB 268 TCACACCTGGGGGTCTGTCACATCGCTTACGCTGCAACAGGGTGCCCGGATGCTGTG 327  
OY 300 AACCTCTGCATCCAGCCAGCCCATCTCTTGGCGGGCATGATGACAGACCTTTCTG 359  
DB 328 AACCTCTGCATCCAGCCAGCCCATCTCTTGGCGGGCATGATGACAGACCTTTCTG 387  
OY 360 TTTTCCACTTTTGTCTGTCACAGATGTCCTCTGTTGGTGTATGTCATGATCTGTAC 419

```
|||||
Db 388 TTTTCCACTTTTGTCTGTACAGAAATGTCCTCTGCTGTGATGTCCTATGATCTGTAC 447
QY 420 GTGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCGTGATCACC 479
Db 448 GTGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCGTGATCACC 507
QY 480 CTGGCGTGACTCTCTGAGACCACTGGAGTCCTTTATCTGATTCATCTGTGTTACTT 539
Db 508 CTGGCGTGACTCTCTGAGACCACTGGAGTCCTTTATCTGATTCATCTGTGTTACTT 567
QY 540 CTACCTTTACCTCTCTGTAGGCCCCAGAAAATTATCATCTTTTGTGAATCTTGGC 599
Db 568 CTACCTTTACCTCTCTGTAGGCCCCAGAAAATTATCAC-TTTTGTGAAATCTTGGC 626
QY 600 TGTTCCTCAACTTGCTCTGTGACATCCCATCATGAGACATGCTTGGCCGAGC 659
Db 627 TGTTCCTCAACTTGCTCTGTGACATCCCATCATGAGACATGCTTGGCCGAGC 686
QY 660 AATTTCGGGCTGGTGGACCCCTGTGTCACAATTTGATGTTGATATATGTGCATCCTTG 719
Db 687 AATTTCGGGCTGGTGGACCCCTGTGTCACAATTTGATGTTGATATATGTGCATCCTTG 746
QY 720 TGTATCTCTGATCCATCCATCAAGGAACTCAGAGGAAAGCCCTTCTGCACCTGCTTC 779
Db 747 TGTATCTCTGATCCATCCATCAAGGAACTCAGAGGAAAGCCCTTCCGACCTGCTTC 806
QY 780 CCACCTCTGTGTGATTTGACTCTTTATGCGACAGCCATTTATCATGTGTGACCCAG 839
Db 807 CCACCTCTGTGTGATTTGACTCTTTATGCGACAGCCATTTATCATGTGTGACCCAG 866
QY 840 ATATGGGAACCCCAAGAGAGAGAGAAATATCTCTCTGCTGTTTCAAGCCCTTTAATCC 899
Db 867 ATATGGGAACCCCAAGAGAGAGAGAAATATCTCTCTGCTGTTTCAAGCCCTTTAATCC 926
QY 900 CATGCTCAATCCCTTATCTGTAGCTTTAGGAAGTCAAGAGTGAAGATCTTTGAAGAG 959
Db 927 CATGCTCAATCCCTTATCTGTAGCTTTAGGAAGTCAAGAGTGAAGATCTTTGAAGAG 986
QY 960 AGTGCTGGAGTAGAAGGCTTTATGAAGAATATGAGCATGTGTGACTGACA 1013
Db 987 AGTGCTGGAGTAGAAGGCTTTATGAAGAATATGAGCATGTGTGACTGACA 1040

RESULT 7
US-09-761-288-3
; Sequence 3, Application US/09761288
; Patent No. US20020065405A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Prayaga, Sudhirdas
; APPLICANT: Taupier, Raymond J
; APPLICANT: Mishra, Vishnu
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Li, Li
; TITLE OF INVENTION: No. US20020065405A1el polypeptides and Nucleic Acids Encoding Sam
; FILE REFERENCE: 15966-638
; CURRENT APPLICATION NUMBER: US/09/761,288
; PRIOR APPLICATION NUMBER: 2001-01-16
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/177,839
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/176,134
; PRIOR FILING DATE: 2000-01-14
; PRIOR APPLICATION NUMBER: 60/175,989
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: 60/218,324
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/220,253
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/178,191
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/178,227
```

```
; PRIOR FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: 60/220,590
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1040
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-761-288-3

Query Match          97.0%; Score 983.6; DB 10; Length 1040;
Best Local Similarity 99.4%; Pred. No. 9.2e-311;
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 TAAACACTTCTCTTAAACCATGAGCATTAACCTGATTTCTCTGTGATAGGATATGGG 60
Db 28 TAAACACTTCTCTTAAACCATGAGCATTAACCTGATTTCTCTGTGATAGGATATGGG 87
QY 61 GACAATATTAACATCCATCAAGAGTTCCTCTACTGGGATTTCCCGTTGGCCCAAGATT 120
Db 88 GACAATATTAACATCCATCAAGAGTTCCTCTACTGGGATTTCCCGTTGGCCCAAGATT 147
QY 121 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTGTTAGTCTTCAACCTGCTGGGGAACGG 180
Db 148 CAGATGCTCCTCTTTGGGCTCTTCTCCCTGTGTTAGTCTTCAACCTGCTGGGGAACGG 207
QY 181 ACCATPCTGGGGCTCATCTCATCTGAGCTCCAGACTCCAGCCGCC-TCATCTTCTCTC 239
Db 208 ACCATPCTGGGGCTCATCTCATCTGAGCTCCAGACTCCAGCCGCCCATGATCTTCTC 267
QY 240 TCACACCTGGCGGTGTCGACATCGCCTACGCGCTGCAACACGGTGCCCGGATGCTGTG 299
Db 268 TCACACCTGGCGGTGTCGACATCGCCTACGCGCTGCAACACGGTGCCCGGATGCTGTG 327
QY 300 AACCTCTGATCCAGCCCAAGCCCATCTCTTTGGGGCCGATGATGACAGACCTTTCTG 359
Db 328 AACCTCTGATCCAGCCCAAGCCCATCTCTTTGGGGCCGATGATGACAGACCTTTCTG 387
QY 360 TTTTCCACTTTTGTCTGTACAGAAATGTCCTCTCTCTGCTGTGATGTCCTATGATCTGTAC 419
Db 388 TTTTCCACTTTTGTCTGTACAGAAATGTCCTCTCTCTGCTGTGATGTCCTATGATCTGTAC 447
QY 420 GTGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCGTGATCACC 479
Db 448 GTGGCCATCTGCCACCCCTCCGATATTGGCCATCATGACCTGGAGAGTCGTGATCACC 507
QY 480 CTGGCGTGACTCTCTGAGACCACTGGAGTCCTTTATCTGATTCATCTGTGTTACTT 539
Db 508 CTGGCGTGACTCTCTGAGACCACTGGAGTCCTTTATCTGATTCATCTGTGTTACTT 567
QY 540 CTACCTTTACCTCTCTGTAGGCCCCAGAAAATTATCATCTTTTGTGAATCTTGGC 599
Db 568 CTACCTTTACCTCTCTGTAGGCCCCAGAAAATTATCAC-TTTTGTGAAATCTTGGC 626
QY 600 TGTTCCTCAACTTGCTCTGTGACATCCCATCATGAGACATGCTTGGCCGAGC 659
Db 627 TGTTCCTCAACTTGCTCTGTGACATCCCATCATGAGACATGCTTGGCCGAGC 686
QY 660 AATTTCGGGCTGGTGGACCCCTGTGTCACAATTTGATGTTGATATATGTGCATCCTTG 719
Db 687 AATTTCGGGCTGGTGGACCCCTGTGTCACAATTTGATGTTGATATATGTGCATCCTTG 746
QY 720 TGTATCTCTGATCCATCCATCAAGGAACTCAGAGGAAAGCCCTTGCACCTGCTTTC 779
Db 747 TGTATCTCTGATCCATCCATCAAGGAACTCAGAGGAAAGCCCTTGCACCTGCTTTC 806
QY 780 CCACCTCTGTGTGATTTGACTCTTTATGCGACAGCCATTTATCATGTGTGACCCAG 839
Db 807 CCACCTCTGTGTGATTTGACTCTTTATGCGACAGCCATTTATCATGTGTGACCCAG 866
QY 840 ATATGGGAACCCCAAGAGAGAGAAATATCTCTGCTGTTTCAAGCCCTTTAATCC 899
Db 867 ATATGGGAACCCCAAGAGAGAGAAATATCTCTGCTGTTTCAAGCCCTTTAATCC 926
```



Db 867 ATATGGGAACCCCAAGGAGCAGAGAATATCTCCTGCTGTTTCACAGCCCTTTAATCC 926  
QY 900 CATGCTCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAGTGAAGATACTTTGAAGAG 959  
Db 927 CATGCTCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAGTGAAGATACTTTGAAGAG 986  
QY 960 AGTGTGGAGTAGAAGGGCTTTATGAAAAGGATTAAGCATTTGTGACTGACA 1013  
Db 987 AGTGTGGAGTAGAAGGGCTTTATGAAAAGGATTAAGCATTTGTGACTGACA 1040

## RESULT 8

US-09-761-288-34  
; Sequence 34, Application US/09761288  
; Patent No. US20020065405A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Prayaga, Sudhirdas  
; APPLICANT: Taupler, Raymond J  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Li, Li  
; TITLE OF INVENTION: No. US20020065405A1el Polypeptides and Nucleic Acids Encoding Sam  
; FILE REFERENCE: 15966-638  
; CURRENT APPLICATION NUMBER: US/09/761,288  
; CURRENT FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: 60/177,839  
; PRIOR FILING DATE: 2000-01-25  
; PRIOR APPLICATION NUMBER: 60/176,134  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 60/175,989  
; PRIOR FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: 60/218,324  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/220,253  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: 60/178,191  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/178,227  
; PRIOR FILING DATE: 2000-01-26  
; PRIOR APPLICATION NUMBER: 60/220,590  
; PRIOR FILING DATE: 2000-07-25  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 34  
; LENGTH: 1040  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-761-288-34

Query Match 97.0%; Score 983.6; DB 10; Length 1040;  
Best Local Similarity 99.4%; Pred. No. 9.2e-311;  
Matches 1008; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 1 TAAACACTTCTCCTAAACCATGAGCATTAATGATTCCTCTGTGCATAGGATATGGGG 60  
Db 28 TAAACACTTCTCCTAAACCATGAGCATTAATGATTCCTCTGTGCATAGGATATGGGA 87  
QY 61 GACATATATACATCCATCACAGAGTTCCTCTACTGCGATTCCCGTGGCCCAAGATT 120  
Db 88 GACATATATACATCCATCACAGAGTTCCTCTACTGCGATTCCCGTGGCCCAAGATT 147  
QY 121 CAGATGCTCTCTTTGGGCTCTTCTCCCTGTCTACGTTTCACCCCTGCTGGGGAACGGG 180  
Db 148 CAGATGCTCTCTTTGGGCTCTTCTCCCTGTCTACGTTTCACCCCTGCTGGGGAACGGG 207  
QY 181 ACCATCTGGGCTCATCTCACTGGAGTCCAGACTGCACGCCCCC-TGTACTTCTTCCTC 239  
Db 208 ACCATCTGGGCTCATCTCACTGGAGTCCAGACTGCACGCCCCCATGTACTTCTTCCTC 267  
QY 240 TCACACCTGGGGTGTGTCACATCGCCTACGCGCTGCAACACGGGTGCCCGGATGCTGTG 299

Db 268 TCACACCTGGGGTGTGTCACATCGCCTACGCGCTGCAACACGGGTGCCCGGATGCTGTG 327  
QY 300 AACCTCTGCATCCAGCCCAAGCCCATCTCCTTTGCGGGCCGCATGATGACACCTTCTG 359  
Db 328 AACCTCTGCATCCAGCCCAAGCCCATCTCCTTTGCGGGCCGCATGATGACACCTTCTG 387  
QY 360 TTTTCACCTTTTGTCTGTACAGAAATGTCTCCTCTGCTGTGTGTATGATCTGTAC 419  
Db 388 TTTTCACCTTTTGTCTGTACAGAAATGTCTCCTCTGCTGTGTGTATGATCTGTAC 447  
QY 420 GGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCACC 479  
Db 448 GTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTCTGCATCACC 507  
QY 480 CTGCGGTGACTTCTCTGAGACCACTGAGTCCCTTTATCTTGATTCATCTGTGTTACTT 539  
Db 508 CTGCGGTGACTTCTCTGAGACCACTGAGTCCCTTTATCTTGATTCATCTGTGTTACTT 567  
QY 540 CTACCTTTACCTTCTGTAGGCCCCAGAAATTTATCATCTTTTGTGAAATCTTGGC 599  
Db 568 CTACCTTTACCTTCTGTAGGCCCCAGAAATTTATCAC-TTTTGTGAAATCTTGGC 626  
QY 600 TGTTCGAACTTGTGCTGTGAGATATCCACATCATAGAGAACATGTGCTTGGCCGAGC 659  
Db 627 TGTTCGAACTTGTGCTGTGAGATATCCACATCATAGAGAACATGTGCTTGGCCGAGC 686  
QY 660 AATTTCTGGGCTGTGGGACCTTGTCCACAATTTGATTTTCATATATGTGCATCTCTG 719  
Db 687 AATTTCTGGGCTGTGGGACCTTGTCCACAATTTGATTTTCATATATGTGCATCTCTG 746  
QY 720 TGTATCTTTCAGATTCATCAAGGGAAGTTTCAGAGAAAGCCTTCTGCACCTGCTTCTC 779  
Db 747 TGTATCTTTCAGATTCATCAAGGGAAGTTTCAGAGAAAGCCTTCTGCACCTGCTTCTC 806  
QY 780 CCACCTCTGTGTGATTTGACTCTTTTATGGCACAGCCATTAATGATGTGGAACCCAG 839  
Db 807 CCACCTCTGTGTGATTTGACTCTTTTATGGCACAGCCATTAATGATGTGGAACCCAG 866  
QY 840 ATATGGGAACCCCAAGGAGCAGAGAATATCTCCTCTGCTTTTCACAGCCCTTTAATCC 899  
Db 867 ATATGGGAACCCCAAGGAGCAGAGAATATCTCCTCTGCTTTTCACAGCCCTTTAATCC 926  
QY 900 CATGCTCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAGTGAAGATACTTTGAAGAG 959  
Db 927 CATGCTCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAGTGAAGATACTTTGAAGAG 986  
QY 960 AGTGTGGAGTAGAAGGGCTTTATGAAAAGGATTAAGCATTTGTGACTGACA 1013  
Db 987 AGTGTGGAGTAGAAGGGCTTTATGAAAAGGATTAAGCATTTGTGACTGACA 1040

## RESULT 9

US-09-864-761-13963/C  
; Sequence 13963, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR APPLICATION NUMBER: 2001-05-23  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04



```
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13963
; LENGTH: 1957
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004889.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; US-09-864-761-13963
```

```
Query Match 53.7%; Score 545; DB 10; Length 1957;
Best Local Similarity 74.6%; Pred. No. 3.1e-167;
Matches 711; Conservative 0; Mismatches 240; Indels 2; Gaps 2;
```

```
QY 36 TTTCCTCTGTCATAGGATATGGGGGACATATACATCCATCAGAGTCCCTCTACT 95
DB 1073 TTTCTTTTTCACAGGGAATGGGGAAATCAGACATGGTTCACAGAGTCCCTCTACT 1014
QY 96 GGGATTCCCGTTGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTCTCCCTGTCTA 155
DB 1013 GGGATTCTCTGGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTCTCCCTGTCTA 954
QY 156 CGTCTTCACCTCTGGGGAACGGACCACTACTGGGCTCATCTCAGTGGACTCCAGACT 215
DB 953 TATCTTCACCTCTGGGGAACGGGCACTCTGGGCTCATCTCAGTGGACTCCAGACT 894
QY 216 GCACGCCCCC-TGTACTTCTCTCTCACAACCTGGCGTGTGACACAGCCCTAGCCCTG 274
DB 893 CCACACCCCATGTACTCTCTCTCACAACCTGGCTGTGTCGACATCGCCTACACCCG 834
QY 275 CAACACGGTGCCCGGATGCTGTGAACCTCTGATCCAGCCCAAGCCCATCTCTTGC 334
DB 833 CAACACGGTGCCCGGATGCTGTGAACCTCTGATCCAGCCCAAGCCCATCTCTTGC 774
QY 335 GGGCCGATGATGACAGACCTTTCTGTTTCCACTTTTGTCTGTACAGAAATGCTCTCT 394
DB 773 TGGTTGCATGACGACAGACCTTTCTGTTTGTGAGTTTGTGACACAGCAATGCTCTCT 714
QY 395 GGTGGTGTATGTCATGATCTGTAGGTGGCACTGTCCACCCCTCCGATATTGGCCAT 454
```

```
DB 713 GGTGCTGATGTCTCAGATCGTTACGTGGCCATCTGCCACCTCTCCGATCTCGTCA 654
QY 455 CATGACCTGGAGAGTCTGCATCAACCTCGCGGTGACTTCTCGACCACTGAGTCTT 514
DB 653 CATGACCTGGAGAGTCTGCATCAACCTCGCGGTGACTTCTCGACCACTGAGTCTT 594
QY 515 ATCTTGATTCATCTGTGTACTCTTACCTTACCTTACCTTCTGTAGCCCGAGAAATTTA 574
DB 593 GGTCTGGCCCATGTGTTCTCATCTTAAGACTGCCCTTCTGTGGCCCTCATGAAATCAA 534
QY 575 TCACCTTTTGTGAAATCTGGCTGTCTCAAACTTGCCTGTGACAGATACCCACATCA 634
DB 533 CCAC-TTCTCTGTGAATCTGTCTGTCTCAGGCTGGCTGTGCTGACACCTGCTCA 475
QY 635 ATGAGAACAATGCTTGGCCGAGCAATTTCTGGGCTGTGGGACCCCTGTCCACATTTG 694
DB 474 ACCAGTGTGCAATCTTGCAGCCTGGCTGTCTTCTCTGCTGGGCGCACCCAGCTGTGC 415
QY 695 TAGTTTCAATATGTGATCCTCTGTCTATCTTCAATCCAATCAAGGAGTTTCA 754
DB 414 TTGCTCTCTACATCGACATCTGTGGCGCATCTGAGATCCAGTCTGGGAGGCGCCA 355
QY 755 GGAAGCCTTCTGCACCTGCTTCTCCACCTCTGTGTGATTTGACTCTTTATGGACAG 814
DB 354 GAAAGCCTTCTCCACCTGCTCTCCACCTCTGCGTGTGGACTCTTTTGGCAGTG 295
QY 815 CCATTATCATGATGTTGGACCCAGATATGGGAACCCCAAGGAGCAGAAATATCTCC 874
DB 294 CCATCATCATGATGATGCCCCCAAGTCCCGCCATCTCTGAGGAGCAGCAAAAGTCTTT 235
QY 875 TGCTGTTTACAGCCTCTTAAATCCCATGCTCAATCCCTTATCTGTAGTCTTAGAAGT 934
DB 234 TTTCTATTACAGTTTCTTCAACCCCAACACTTAAACCCCTGATTTACAGCCTGAGGAAG 175
QY 935 CAGAAGTGAAGATCTTTGAAGAGAGTGTGGGAGTAGAAGGCTTTATGA 987
DB 174 GAGAGTCAAGGGTCCCTTGAGGAGAGCAGTGGCCAGGAAGTCAATTCCTAA 122
```

## RESULT 10

```
; US-10-098-841-92
; Sequence 92, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyun
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
```

```
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 92
; LENGTH: 2282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1257)..(2189)
US-10-098-841-92
```

```
Query Match      53.1%; Score 538.2; DB 9; Length 2282;
Best Local Similarity 73.6%; Pred. No. 5.6e-165;
Matches 712; Conservative 0; Mismatches 253; Indels 2; Gaps 2;

QY 31 CTGTGATTTCCCTCTGTCATAGGATATGGGGGACATATACATCCATCCACAGAGTCTC 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1233 CTTGTGTTGTTTTAGTAGTGAATGSGTGAATAATCAGACATGCTCACAGAGTCTC 1292

QY 91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTCTCCCTG 150
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1293 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTCTCCCTG 1352

QY 151 TTCTAGCTCTTCACCCCTGCTGGGGAAGGGGACCATCTACTGGGCTCATCTCAGTGGACTCC 210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1353 TTCTATGCTCTACCCCTGCTGGGGAATGGGACCATCTCTGGGCTCATCTCAGTGGACTCC 1412

QY 211 AGACTGCACGCCCC-TGTACTTCTCTCTCACACCTGGCGGTGCTGCACATCGCCTAC 269
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1413 AGACTGCACACCCCATGTACTTCTCTCTCACACCTGGCGGTGCTGCACATCGCCTAT 1472

QY 270 GCCTGCACACAGGTGCCCCGGATGCTGTGAACCTCTGCATCCAGCCAGCCCATCTCC 329
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1473 GCCTGCACACAGTGCCTCCAGATGCTGTGAACCTCTGCATCCAGCCAGCCCATCTCC 1532

QY 330 TTTCGGGGCGGATGATGACAGACCTTCTGTTTCCACTTTTGTCTGTACAGAAATGCTC 389
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1533 TTTGCTGGCTGCATGACACAGACCTTCTCTTTTGAAGTTTGACATCTGAATGCTC 1592

QY 390 CTCCTGGTGGATGTCCTATGATCTGACGTGGCCATCTGCCACCCCTCCGATATTG 449
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1593 CTGTTGGTGTGATGTCCTACGATCGTACGTGGCCATCTGCCACCCCTCTCCGATATTG 1652

QY 450 GCCATCATGACCTGAGAGTCTGCATCACCCCTCGCGGTGACTTCTTGAGCCACTGAGTTC 509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1653 ATCATCATGACCTGAGAAAGTCTGCATCCTGCGCATCACTTCTTGACATGTGGCTCC 1712

QY 510 CTTTATCCTGATCATCTTGTGTACTTCTACCTTTACCTTCTGTAGCCCCAGAAA 569
    || || || || || || || || || || || || || || || || || || || || || || ||
DB 1713 CTCCTGGCTATGATCCATGTGAGCCTCATCTTAAGACTGCCCTTTGTGGCCTCGTGA 1772

QY 570 AATTATCACTTTTTTTGGAATCTTGGCTGTCTCAAACTTGCCTGTGACAGATACCA 629
    || || || || || || || || || || || || || || || || || || || || || || ||
DB 1773 ATCAACACAC-TTCTTCTGTGAATCTCTGTCTGCTCAGGCTGGCCTGTGATACCTG 1831

QY 630 CATCAATGAGAATGCTCTGGCCGAGCAATTTCTGGGCTGTGGGACCTTGTCCAC 689
    ||||| || || || || || || || || || || || || || || || || || || || || ||
DB 1832 GCTCAACGAGGTGCTCATCTTTCAGCCCTGCATGTCATCTCTGAGGACCACTCTGCT 1891

QY 690 AATTGTAGTTTCATATATGTGCATCCTCTGTGCTATCTCTCAGATCCAATCAAGGAGT 749
    || || || || || || || || || || || || || || || || || || || || || || ||
DB 1892 GGTGCTGCTCTCTACTACACATCTCTGGCCGACCATCTTGAGGATCCAGTCTGGGAGGG 1951

QY 750 TCAGAGGAAAGCCTCTGCACCTGCTTCTCCACCTCTGTGTGATTTGACTCTTTATGG 809
    || || || || || || || || || || || || || || || || || || || || || || ||
DB 1952 CCGCAGAAAGGCTTCTCCACCTGCTCTCCACCTCTGCGTAGTGGGACTCTTCTTTGG 2011

QY 810 CACAGCCATTAATCATGATATGTGGACCCAGATATGGGAACCCCAAGSAGCAGAAATA 869
    || ||||| ||||| ||||| || || || || || || || || || || || || || || || ||
DB 2012 CAGCGCCATGTCATGTACATGCCCCCTAAGTCCGCCCATCTGAGAGCAGCAGAAAGT 2071

QY 870 TCTCCTGCTGTTTCACAGCCTCTTTAATCCCATGCTCAATCCCTTATCTGTAGTCTTAG 929
    || || || || || || || || || || || || || || || || || || || || || || ||
DB 2072 CCTTTTCTATTATTACAGTCTTTTCACACCCGATGCTAAACCCCTGATTTACACCTGAG 2131
```

```
QY 930 GAACTCAGAGTGAAGATTTTGAAGAGAGTCTGGGAGTGAAGAGGCTTTATGAAA 989
    ||| ||| ||| ||| ||| ||||| ||| ||| ||||| ||| |||
DB 2132 GAATGTAGGCTCAAGGCTGCTGAGAGAGACACTGTGCAAGAAAGTCAATCTTAAGA 2191

QY 990 AGGATTA 996
    | | |
DB 2192 GGTGTGA 2198
```

```
RESULT 11
US-09-747-835A-62
; Sequence 62, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-
; TITLE OF INVENTION: (LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747, 835A
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729, 739
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: US 09/653, 450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620, 312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598, 042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552, 317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488, 725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 62
; LENGTH: 2282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1257)..(2189)
US-09-747-835A-62
```

```
Query Match      53.1%; Score 538.2; DB 10; Length 2282;
Best Local Similarity 73.6%; Pred. No. 5.6e-165;
Matches 712; Conservative 0; Mismatches 253; Indels 2; Gaps 2;

QY 31 CTGTGATTTCCCTCTGTCATAGGATATGGGGGACATATACATCCATCCACAGAGTCTC 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1233 CTTGTGTTGTTTTAGTAGTGAATGSGTGAATAATCAGACATGCTCACAGAGTCTC 1292

QY 91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTCTCCCTG 150
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1293 CTACTGGGATTTCTCCTGGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTCTCCCTG 1352

QY 151 TTCTAGCTCTTCACCCCTGCTGGGGAAGGGGACCATCTACTGGGCTCATCTCAGTGGACTCC 210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1353 TTCTATGCTCTTCACCCCTGCTGGGGAATGGGACCATCTCTGGGCTCATCTCAGTGGACTCC 1412

QY 211 AGACTGCACGCCCC-TGTACTTCTCTCTCACACCTGGCGGTGCTGCACATCGCCTAC 269
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1413 AGACTGCACACCCCATGTACTTCTCTCTCTCACACCTGGCGGTGCTGCACATCGCCTAT 1472

QY 270 GCCTGCACACAGGTGCCCCGGATGCTGTGAACCTCTCTGATCCAGCCAGCCATCTCC 329
```

```

1473 GCGTCACACAGTGGCCGAGATGCTGTGAACCTCTGTCATCCAGCCCAAGCCCAATCTCC 1532
QY 330 TTGGGGGCGGCATGATGACAGCTTCTGTCTTCCACTTTTGGCTGTACAGAAATGCTC 389
Db 1533 TTGGCTGGCTGATGACACAGACCTTCTCTTTTGTAGTTTGCACATACTGAATGCTC 1592
QY 390 CTCTGGTGGTGTATGTCCTATGATCTGTACGTTGGCCATCTGCCACCCCTCCGATATTG 449
Db 1593 CTGTGGTGGTGTATGTCCTATGATCTGTACGTTGGCCATCTGCCACCCCTCCGATATTG 1652
QY 450 GCCATCATGACCTGAGAGTCTGCATCAACCCTGCGGTGACTTCTGTGACCATGAGTTC 509
Db 1653 ATCATCATGACCTGAGAAAGTCTGCATCACTCTGCGCATCTTCTGTGACCATGAGTTC 1712
QY 510 CTTTATCCTTGATTCATCTTGTGTAATCTTCTTCTTACCTTTTACCTTCTGTGAGCCCAAAA 569
Db 1713 CTCTGGCTATGCTCATATGATGAGCCCTCATCTTAAGACTGCCCTTTTGTGGGCTCTGTA 1772
QY 570 ATTATACATTTTCTTGTGTAATCTTGGCTGTCTCAAACTTGCCTGTGACATACCA 629
Db 1773 ATCAACCAC-TTCTTCTGTGAATCTTGTCTCTCAAGGCTGCTGTGTATACCTG 1831
QY 630 CATCAATGAGAACATGCTCTTGGCCGAGCAATTTCTGGCTGTGGGACCTTGTCCAC 689
Db 1832 GCTCAACCGAGTGTATCTTGTGACGCTGCATGTTCATCTGTGGGACCACTGTGCT 1891
QY 690 AATGTAGTTTCAATATATGTGCATCTCTGTCTATCTTCAATCAAGGAAGT 749
Db 1892 GGTGCTGTCTCTACTACACATCTCTGGCCATCTCTGAGGATCCAGTCTGGGAGGG 1951
QY 750 TCAGAGAAAGCCTTCTGCACCTGCTTCCACCTCTGTGTGATGTGACTCTTTTATGG 809
Db 1952 CCGCAGAAAGCCTTCTCTCCACCTGCTCTCCACCTCTGTGTGAGTGTGACTCTTTTGG 2011
QY 810 CACAGCCATATATATGTATGTGAGCCAGATATGGAACCCCAAGAGCAGAGAAATA 869
Db 2012 CAGCGCCATCTCATGTATGAGGCCCCCTTAAGTCCCGCATCTGAGGAGCAGAGAAAGT 2071
QY 870 TCTCTGTCTGTTCACAGCCTCTTAAATCCCATGCTCAATCCCTTATCTGTAGTCTAG 929
Db 2072 CCTTTTCTATTTTACAGTCTCTTCAACCCGATGCTTAACCCCTGATTTACAACCTGAG 2131
QY 930 GAACCTCAGAGTGAAGATCTTGAAGAGAGTCTGGAGTAGAAGGCTTTATGAA 989
Db 2132 GAATGTAGAGTCAAGGGTGCCTGAGGAGAGCACTGTGCAAGGAAGTCAATCTCTAAGA 2191
QY 990 AGGATTA 996
Db 2192 GGTGTGA 2198

RESULT 12
US-09-747-835A-35
; Sequence 35, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Zhang, Dunrui
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Dmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE
; TITLE OF INVENTION: (LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
```

```

; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; SOFTWARE: Patent version 3.0
; SEQ ID NO 35
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (833)..(1417)
; US-09-747-835A-35

Query Match          52.9%; Score 536.6; DB 10; Length 1788;
Best Local Similarity 73.5%; Pred. No. 1.6e-164;
Matches 711; Conservative 0; Mismatches 254; Indels 2; Gaps 2;

QY 31 CTGATTTCTCTGTATAGGATATGGGGACATATACATCCATCAGAGTTCCTC 90
Db 461 CTTTGTGTTTGTATAGTGAATGTTGAATAATCAGACAATGTTACAGAGTTCCTC 520
QY 91 CTACTGGATTTCCCGTGGCCCAAGGATTCAGATGCTCTTGGGCTCTCTCCCTG 150
Db 521 CTACTGGATTTCTCTCGGCCCCAAGGATTCAGATGCTCTTGGGCTCTCTCCCTG 580
QY 151 TTCTACGCTCTACCCCTGCTGGGGAACGGACCATCTGGGCTCATCTCACTGACTCC 210
Db 581 TTCTATGCTTACCCCTGCTGGGGAATGGACCATCTGGGCTCATCTCACTGACTCC 640
QY 211 AGACTGCACGCCCC-TGTACTTCTCTCTCTACACCTGGCGTGTGACATGCTCTAC 269
Db 641 AGACTGCACACCCCATGTACTTCTCTCTCTACACCTGGCGTGTGACATGCTCTAT 700
QY 270 GCGTGAACACAGGTGCCCCGATGCTGTGAACTCTCTGATCCAGCCAGCCATCTCC 329
Db 701 GCGTGAACACAGGTGCCCCGATGCTGTGAACTCTCTGATCCAGCCAGCCATCTCC 760
QY 330 TTGGGGGCGCATGATGACAGCCTTCTGTGTTTCCACTTTTGTCTGTCACAGATGCTC 389
Db 761 TTGGCTGGCTGCATGACATAGACCTTCTCTTTTGTGATTTGACATAGTATGCTC 820
QY 390 CTCTGGTGGTGTATGCTCTATGATCTGTACGTTGGCCATCTGCCACCCCTCCGATATTG 449
Db 821 CTGTGGTGGTGTATGCTCTATGATCTGTACGTTGGCCATCTGCCACCCCTCCGATATTG 880
QY 450 GCCATCATGACCTGAGAGTCTGCATCAACCCTGCGGTGACTTCTTGACCATGAGTTC 509
Db 881 ATCATCATGACCTGAGAAAGTCTGCATCACTCTGCGCATCACTTCTTGACATGAGTTC 940
QY 510 CTTTATCCTTGATTCATCTTGTGTAATCTTCTTACCTTTTACCTTCTGTAGCCCAAAA 569
Db 941 CTCTGGCTATGCTCATGTGAGCCTTCACTTAAGACTGCCCCCTTTTGTGGCCCTGTA 1000
QY 570 ATTATACATTTTCTTGTGTAATCTTGGCTGTCTCAAACTTGCCTGTGACATACCA 629
Db 1001 ATCAACCAC-TTCTTCTGTGAATCTTGTCTCTCAGGCTGGCCCTGTGTATACCTG 1059
QY 630 CATCAATGAGAACATGCTCTTGGCCGAGCAATTTCTGGGCTGTGGGACCTTGTCCAC 689
Db 1060 GCTCAACCGAGTGTATCTTGTGAGCCTGCATGTTCATCTCTGTGGGACCACTGTGCT 1119
QY 690 AATGTAGTTTCAATATATGTGCATCTCTGTGCTATCTTCAATCAAGGAAGT 749
Db 1120 GGTGCTGTCTCTACTACACATCTCTGGGCGCATCTGAGGATCCAGTCTGGGAGGG 1179
```

QY 750 TCAGAGAAAGCCTTCTGCACCTGCTCTCCACCCTGTGTGATGTGACTCTTTATGG 809  
Db 1180 CCGAGAAAGGCTTCTCCACCTGCTCCCTCCACCCTGTGGTAGTGGACTCTTTTGG 1239  
QY 810 CACAGCCATTCATGATGTATGTGACCCAGATATGGAACCCCAAGAGCAGAGAATA 869  
Db 1240 CAGCGCCATGCTATGATGAGCCCTAAGTCCCGCCATCTGAGGAGCAGAGAAGT 1299  
QY 870 TCTCCTGCTGTTCACAGCCTCTTAAATCCCATGCTCAATCCCTTATCTGTAGTCTAG 929  
Db 1300 CCTTTTCTATTTCAGTCTTTCACCCGATGCTAAACCCCTGATTTCACACCTGAG 1359  
QY 930 GAACCTCAGAGTGAAGATACTTTGAGAGAGAGTGTGGAGTAGAAGGCTTTATGAAA 989  
Db 1360 GAATGTAGAGTCAAGGGTGCCCTGAGAGAGACACTGTGCAAGGAAGTCAATTCCTAAGA 1419  
QY 990 AGGATTA 996  
Db 1420 GGTGTGA 1426

## RESULT 13

US-09-747-835A-34  
; Sequence 34, Application US/09747835A  
; Patent No. US20020146692A1  
; GENERAL INFORMATION:  
; APPLICANT: Yamazaki, Victoria  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Zhou, Ping  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyun  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Drmanac, Radoje T  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-LIKE  
; TITLE OF INVENTION: LIKE POLYPEPTIDES AND POLYNUCLEOTIDES  
; FILE REFERENCE: HVS-37CIP  
; CURRENT APPLICATION NUMBER: US/09/747, 835A  
; PRIOR APPLICATION NUMBER: 2002-03-08  
; PRIOR FILING DATE: US 09/729, 739  
; PRIOR FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: US 09/653, 450  
; PRIOR FILING DATE: 2000-08-31  
; PRIOR APPLICATION NUMBER: US 09/620, 312  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: US 09/598, 042  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: US 09/552, 317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: US 09/488, 725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 34  
; LENGTH: 2735  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-747-835A-34

Query Match 52.9%; Score 536.6; DB 10; Length 2735;  
Best Local Similarity 73.5%; Pred. No. 2.1e-164;  
Matches 711; Conservative 0; Mismatches 254; Indels 2; Gaps 2;

QY 31 CTGATTTCTCTGTGATAGGATATGGGACATATATACATCCATCAGAGTTCCTG 90  
Db 1311 CTTGTTGTTTATAGTAGAATGCTGAAAAATCAGACAAATGCTCAGAGTTCCTG 1370  
QY 91 CTACTGGATTTCCCTGGCCCAAGATTGATGCTCTCTTGGGCTCTTCCCTG 150  
Db 1371 CTACTGGATTTCTCTGGGCCCAAGATTGATGCTCTCTTGGGCTCTTCCCTG 1430  
QY 151 TTCTACGTCTTCAACCTGCTGGGAACGGAGACATACCTGCGCTCATCTCACTGACTCC 210

Db 1431 TTTATGTTCTTCAACCTGCTGGGAATGGAGACATCTGGGGCTCATCTGACTGCC 1490  
QY 211 AGACTGACGCCCCC-TGTACTTCTTCTCTCACACCTGGCGGTGCTGACATCGCTAC 269  
Db 1491 AGACTGACACCCCATGTACTTCTTCTCTCACACCTGGCGGTGCTGACATCGCTAT 1550  
QY 270 GCTGCAACACGGTGGCCCGGATGCTGTGAACCTCTGATCCAGCCCAAGCCATCTCC 329  
Db 1551 GCTGCAACACAGTGGCCCGGATGCTGTGAACCTCTGATCCAGCCCAAGCCATCTCC 1610  
QY 330 TTGCGGGCCGATGATGACAGCCTTCTGTTTCCACTTTTGTGCTGACAGAAATGCTC 389  
Db 1611 TTGCTGCTGATGATGACATGACCTTCTCTTTTGAATTTGACATGATGATGCTC 1670  
QY 390 CTCTGCTGCTGATGCTCTGATGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 449  
Db 1671 CTGCTGCTGCTGATGCTCTGATGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 1730  
QY 450 GCCATCATGACCTGAGAGTGTGATCACCCTCGCGGTGACTTCTGACACCTGAGTCC 509  
Db 1731 ATCATCATGACCTGAGAGTGTGATCACCCTCGCGGTGACTTCTGACACCTGAGTCC 1790  
QY 510 CTTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 569  
Db 1791 CTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1850  
QY 570 ATTATGACTTTTCTTGTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 629  
Db 1851 ATCAACAC-TCCTCTGTGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1909  
QY 630 CATCAATGACATGATGCTGTGCGCGGAGCAATTTCTGGCTGTGGACCTGTGCCAC 689  
Db 1910 GCTCAACCAAGTGTGATCTTGTGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1969  
QY 690 AATGTAGTTTCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 749  
Db 1970 GGTGCTGCTGCTTACTACACATCTGCGGGCCATCTGAGGATCCAGTCTGGGAGGG 2029  
QY 750 TCAGAGAAAGCCTTCTGCACCTGCTCTCCACCCTGTGTGATGTGACTCTTTATGG 809  
Db 2030 CCGAGAAAGGCTTCTCCACCTGCTCCCTCCACCCTGTGTGATGTGACTCTTTATGG 2089  
QY 810 CACAGCCATTCATGATGTATGTGACCCAGATATGGAACCCCAAGAGCAGAGAATA 869  
Db 2090 CAGCGCCATGCTATGATGAGCCCTAAGTCCCGCCATCTGAGGAGCAGAGAAGT 2149  
QY 870 TCTCCTGCTGTTCACAGCCTCTTAAATCCCATGCTCAATCCCTTATCTGTAGTCTAG 929  
Db 2150 CCTTTTCTATTTCAGTCTTTCACCCGATGCTAAACCCCTGATTTCACACCTGAG 2209  
QY 930 GAACCTCAGAGTGAAGATACTTTGAAGAGAGTGTGGAGTAGAAGGCTTTATGAAA 989  
Db 2210 GAATGTAGAGTCAAGGGTGCCCTGAGAGAGACACTGTGCAAGGAAGTCAATTCCTAAGA 2269  
QY 990 AGGATTA 996  
Db 2270 GGTGTGA 2276

## RESULT 14

US-09-864-761-30527/C  
; Sequence 30527, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Ranzel, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecmica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864, 761



```
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30527
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004889.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.93
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: SWISSPROT HIT: P34984, EVALUE 1.00e-86
; OTHER INFORMATION: EST HUMAN HIT: BF116115.1, EVALUE 0.00e+00
; OTHER INFORMATION: NT HIT: U86281.1, EVALUE 0.00e+00
; US-09-864-761-30527
```

```
Query Match 52.8%; Score 535.2; DB 10; Length 933;
Best Local Similarity 74.8%; Pred. No. 3.1e-164;
Matches 697; Conservative 0; Mismatches 233; Indels 2; Gaps 2;
```

```
QY 52 GATATGGGAGACAATATACATCCATCAGAGTTCCTCTACTGGGATTCCCGTTGGC 111
Db 933 GAAATGGGGGAAATACAGACAATGTACAGAGTTCCTCTACTGGGATTTCCTGGGC 874
QY 112 CCAAGATTGAGATGCTCTCTTTGGGCTCTTCCTCTTCTACGTTTCACCCCTGCTG 171
Db 873 CCAAGATTGAGATGCTCTCTTTGGGCTCTTCCTCTTCTATATCTTACACCCCTGCTG 814
QY 172 GGAACGGGACCATATCTGGGCTCATCTCACTGACTGCAGACTGCAGCCGCC-TGTAC 230
Db 813 GGAACGGGACCATCTGGGCTCATCTCACTGACTGCAGACTGCAGCCCATGTAC 754
QY 231 TTCTTCTCTACACCTGCGGCTCGTGCAGATCGGCTAGCCCTGCAACAGGTGCCCGG 290
```

```
Db 753 TTCTTCTCTACACCTGGCTGTCTGTCAGATCGCTTACACCCGCAACAGGTGCCCGG 694
QY 291 ATGCTGGTGAACCTCTCTGCATCCAGCCAAAGCCCATCTCTTTGGGGCCGATGATGAG 350
Db 693 ATGCTGGGCAACCTCTCTGCATCCAGCCAAAGCCCATCTCTTTGGGGCTGATGATGAG 634
QY 351 ACCTTCTGTTTCCACTTTTGTCTGTCAAGATGTCTCTCTGCTGGTGGTATGCTCTAT 410
Db 633 ACCTTCTGTTTGAAGTTTGGACACAGCAATGTCTCTCTGCTGGTGGTATGCTCTAT 574
QY 411 GATCTGACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGAGAGTC 470
Db 573 GATCGTTACGTGGCCATCTGCCACCCCTCCGATATCTCTCTGCTGGTGGTATGCTCTAT 514
QY 471 TGCATCACCCTGGGCTGACTTCTCTGACCACTGAGTCCCTTTATGCTTGAATTCATCTT 530
Db 513 TGCATCACCCTGGGCTGACTTCTCTGACCACTGAGTCCCTCTCTGCTGGTGGTATGCTCTAT 454
QY 531 GTGTTACTTCTACCTTTACCCCTCTGTAGGCCCCAGAAATTTATACCTTTTGTGGA 590
Db 453 GTTCTCATCTTAAGACTGCCCTTCTCTGGGCTCATGAAATCAACAC-TCCTCTGTGA 395
QY 591 AATCTTGCTGTCTCAAACTTGCTGTGACAGATACCCATCAATGAGAACATGCTCTT 650
Db 394 AATCTGTCTGTCTCAAGCTTGCTGTGACAGATACCCATCAATGAGAACATGCTCTT 335
QY 651 GCGCCGAGCAATTTCTGGGCTGTGGACCTTGTCACAAATTTAGTTTATATATGTG 710
Db 334 TGCAGCCTGCTGTCTTCTGCTGTGGGCCCCAGCCAGCCTGTGCTGTCTTCTACTGCA 275
QY 711 CATCTCTGTCTATCTTCAAGATCCAAATCAAGGAGTTCAAGAGAAAGCTTCTGCAC 770
Db 274 CATCTGCGGCTATCTTCAAGATCCAAATCAAGGAGTTCAAGAGAAAGCTTCTGCAC 215
QY 771 CTGCTTCTCCACCTCTGTGTGATTTGACTCTTTATGGACAGCCATATCATGTATGT 830
Db 214 CTGCTTCTCCACCTCTGTGTGATTTGACTCTTTATGGACAGCCATATCATGTATCAT 155
QY 831 TGAACCCAGATATGGAACCCCAAGAGAGAGAGAGAAATATCTCTGCTGTTTACAGCCT 890
Db 154 GCGCCCAAGTCCCGCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 95
QY 891 CTTTATCCCATGCTCAATCCCTTATCTGTAGTCTTATGGAACCTCAGAGTGAAGATAC 950
Db 94 TTTCAACCCCAACACTTAACCCCTGATTTACAGCCTGAGAGAGAGAGAGAGAGAGAG 35
QY 951 TTTGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 982
Db 34 CCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3
```

```
RESULT 15
US-09-747-835A-41
; Sequence 41, Application US/09747835A
; Patent No. US20020146692A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO G PROTEIN-COUPLED RECEPTOR-1.
; TITLE OF INVENTION: (LIKE) POLYPEPTIDES AND POLYNUCLEOTIDES
; FILE REFERENCE: HVS-37CIP
; CURRENT APPLICATION NUMBER: US/09/747,835A
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 09/729,739
; PRIOR FILING DATE: 2000-12-04
```



```
; PRIOR APPLICATION NUMBER: US 09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (485)..(1411)
US-09-747-835A-41
```

```
Query Match 50.4%; Score 511.4; DB 10; Length 1782;
Best Local Similarity 72.7%; Pred. No. 2.8e-156;
Matches 703; Conservative 0; Mismatches 256; Indels 8; Gaps 3;
```

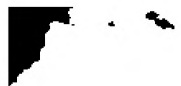
```
OY 31 CTGATTCCTCTGTCATAGGATATGGGACATATTAACATCCATCAGAGTTCCCTC 90
Db 461 CTTGTTGTTTATTAGTAGAATGGTGAATAATCAGACAATGTCACAGAGTTCCCTC 520
OY 91 CTACTGGGATTTCCCGTTGGCCCAAGGATTCAGATGCTCCTTTGGGCTCTCCCTG 150
Db 521 CTACTGGGATTTCTCTGGGCCAAGGATTCAGATGCTCCTTTGGGCTCTCCCTG 580
OY 151 TTCTACGCTTCCACCTGCTGGGGAAGGACCATCTAGGCTCATCTCAGTGGACTCC 210
Db 581 TTCTATATCTTCAACCTGCTGGGGAAGGAGGACCATCTAGGCTCATCTCAGTGGACTCC 640
OY 211 AGACTGCACGCCCC-TGTACTTCTTCTCTCAGACCTGGGGTCTGTCAGATCGCCTAC 269
Db 641 AGACTGCCATACCCCATGTTCTTCTCTCAGACCTGGGGTCTGTCAGATCGCCTAC 700
OY 270 GCCTGCAACACGGTGCCCCGATGCTGTGAACCTCTGTCATCCAGCCAGCCCATCTCC 329
Db 701 ACCCGCAACACGGTGCCCCGATGCTGTGAACCTCTGTCATCCAGCCAGCCCATCTCC 760
OY 330 TTGCGGGCCGATGATGACAGACCTTCTGTTCACATTTTGTCTGTCACAGATGCTCTC 389
Db 761 TTGCGGTGTCATGACGACAGACCTTCTGTGTGAGTTTGGACACAGCGAATGCTCTC 820
OY 390 CTCTGTGTGATGTCCTATGATCTGTAGTGGCCATCTGCCACCCCTCCGATATTG 449
Db 821 CTGCTGTGCTGATGTCCTATGATCTGTAGTGGCCATCTGCCACCCCTCTCCGATATTG 880
OY 450 GCCATCATGACCTGGAGAGTTCGATCACCCTCGCGGTGACTTCTTGACCCACTGGAGTC 509
Db 881 GTCATCATGACCTG-----CTGCATCACTCTGGCCATCACTTCTTGACATGTGCTCC 934
OY 510 CTTTATCCTTGATTCATCTGTGTACTTCTTACCTTTACCCCTTCTGTAGGCCCGAGAAA 569
Db 935 CTCTGTGCTATGTCATGTCAGCCTCATCTTAAGACTGCCCCCTTTGTGGGCTCTGTGAA 994
OY 570 ATTTATCACTTTTGTGAAATCTTGCGTGTCTCAAACTTGCTGTGCAATACCCA 629
Db 995 ATCAACAC-TTCTTCTGTGAATCTGTCTGTCTCAAGGCTGGCCCTGTGCTGATACCTG 1053
OY 630 CATCAATGAGACATGTCCTTGCGCGGAGCAATTTCTGCGCTGGTGAGCCCTGTGTCAC 689
Db 1054 GCTCAACGAGGTGTCATCTTTCAGCCCTGCATGTTCAATCCTGTGGGACCACTGTGCT 1113
OY 690 AATTGTAGTTTCATATATGTCATCTCTGTCTATCTTCAGATCCCAATCAAGGGAAGT 749
Db 1114 GGTGCTGTCTCTCTACTCAGACATCTCTGGGCGCATCTGAGGATCCAGTCTGGGGAAGG 1173
```

```
OY 750 TCAGAGAAAGCCTTCTGCAACCTGCTTCTCCACCTCTGTGATGACTCTTTATGG 809
Db 1174 CCGAGAAAGCCTTCTGCAACCTGCTTCTCCACCTCTGTGATGACTCTTTATGG 1233
OY 810 CACAGCATTTATCATGTAATGAGCCAGATATGGAACCCCAAGGACAGAAATA 869
Db 1234 CAGGECATCTCATATGATGAGCCCTAAGTCCGCACTCTGAGAGACAGAAAGT 1293
OY 870 TCTCTGCTGTTTACAGCCCTTTAATCCCATGCTCAATCCCTTATCTGTAGTCTAG 929
Db 1294 CCTTTCTATTTTACAGTCTTCAACCCGATGCTAAACCCCTGATTTACAACCTGAG 1353
OY 930 GAATCAGAGTGAAGATCTTTGAAGAGAGTGTGGAGTAGAAGGCTTTATGAAA 989
Db 1354 GAATGTAAGGTCAAGGCTGCCCCGAGAGAGACACTGTGCAAGAAAGTATCTCTAAGA 1413
OY 990 AGGATTA 996
Db 1414 GGTGTGA 1420
```

```
Search completed: May 2, 2003, 06:42:32
Job time : 1681 secs
```







128 ATGTGAAAATCAGACATGCTCAGAGCTTCCTCTACTGGGATTTCT 177  
17 oValGlyProArgIleGlnMetLeuPheGlyLeuPheSerLeuPhe 34  
:::|||||  
178 CTTGGGCCCAAGGATTGAGATGCTCTCTTTGGGCTCTCTCCCTGTTCT 227  
34 yValPheThrLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
|||||  
228 ATGCTTCAACCTGCTGGGGAATGGACCATCTGGGGCTCATCTCACTG 277  
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAla 67  
|||||  
278 GACTCCAGACTCCACACCCCATGTAATCTCTCTCTCACACCTGGCCGT 327  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsn 84  
|||||  
328 CGTCAACATGCGCTATGCTGCAACACAGATGCCCATGCTGTGAACC 377  
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
|||||  
378 TCCGTCATCCAGCCCAAGCCCATCTCTTGTGCTGTCATGACACTAGAC 427  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
|||||  
428 TTTCTCTTTTGTAGTTTTCACATAGTGAATGCTCTGTTGCTGCTGAT 477  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAla 134  
|||||  
478 GTCTACGATCGGTACGTGCGCATCTGCCACCCCTCCGATATTCTATCA 527  
134 lMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150  
|||||  
528 TCATGACCTGGAAGTCTGCACTCTGGGCACTCTCTGACATGT 577  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPh 167  
|||||  
578 GGCTCCCTCTGGCTATGCTGTCATGTGAGCCTCATCTTAAGACTGCCCTT 627  
167 eCysArgProGlnLysIleTyrHisPhePheCysGlyIleLeuAlaVal 184  
|||||  
628 TTGTGGGCTCGTGAATCAACCACTCTCTGTGAATCTCTGCTCTCC 677  
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200  
|||||  
678 TCAGGCTGGCTGCTGCTGATACCTGCTCAACCAAGTGTGCTTTGAA 727  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217  
|||||  
728 GCCTGATGTTCACTCTGTTGGGACCACTCTGCTGCTGCTCTCTTA 777  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGln 234  
|||||  
778 CTCACACATCTCTGGGGGCGCATCTGAGGATCCAGTCTGGGAGGGCCGCA 827  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
|||||  
828 GAAAGGCTTCTCCACCTGCTCTCCCACTCTGCTAGTGGGACCTCTTC 877  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProly 267  
:::|||||  
878 TTTGGAGCGCATGCTATGATGAGCCCTAAGTCCCGCATCTCTGA 927  
267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnPro... 282  
:::|||||  
928 GGAGCAGCAGAGGCTCTTTTCTTATTTTACAGTCTCTTCAACCCGCA 977  
283 MetLeuAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsn 298  
|||||  
978 TGCTTAACCCCTGATTTTACACCCCTGAGGAATGTAGAGGTCACAGGT 1027  
299 ThrLeuLys 301  
:::|||||  
1028 GCCCTCCGA 1036

seq\_name: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-08-467-947A-1  
seq\_documentation\_block:  
Sequence 1, Application US/08467947A  
Patent No. 6090575  
GENERAL INFORMATION:  
APPLICANT: LI, YI  
APPLICANT: CAO, LIANG  
APPLICANT: NI, JIAN  
APPLICANT: GENTZ, REINER  
APPLICANT: BULT, CAROL J.  
APPLICANT: SUTTON III, GRANGER G.  
APPLICANT: ROSEN, CRAIG A.  
TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein  
TITLE OF INVENTION: Coupled Receptor GPR1  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,947A  
FILING DATE: 08-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04079  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: STEFFE, ERIC R.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1713 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 116..1003  
US-08-467-947A-1  
alignment\_scores:  
Quality: 1048.50 Length: 305  
Ratio: 4.017 Gaps: 4  
Percent Similarity: 85.574 Percent Identity: 69.180  
alignment\_block:  
US-09-898-586-24 x US-08-467-947A-1 ..  
Align seg 1/1 to: US-08-467-947A-1 from: 1 to: 1713  
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhe 17  
||| ||| ::|||  
128 ATGTGAAAATCAGACATGCTCAGAGCTTCCTCTACTGGGATTTCT 177  
||| ||| |||  
17 oValGlyProArgIleGlnMetLeuPheGlyLeuPheSerLeuPhe 34  
:::|||||  
178 CTTGGGCCCAAGGATTGAGATGCTCTCTTTGGGCTCTCTCCCTGTTCT 227



